



## 506: Establishing & Maintaining an Effective Best Employment Practices Audit Program

**Allan G. King**  
*Attorney*  
Littler Mendelson

**Mark Schwartz**  
*Vice President & Assistant General Counsel, Labor and Employment*  
Waste Management, Inc.

**Bernard R. Siskin, Ph.D.**  
*Managing Director*  
LECG, LLC

## Faculty Biographies

### Allan G. King

Allan G. King is a shareholder in Littler Mendelson's Dallas office. He is board certified in employment and labor law by the Texas Board of Legal Specialization. He specializes in issues relating to the statistical proof and defense of employment discrimination.

Prior to becoming an attorney, Mr. King was a member of the economics department faculty of the University of Texas at Austin, specializing in labor economics and demography.

Mr. King received a BA from the City College of New York, a master's and Ph.D. from Cornell University, and is a graduate of the University of Texas School of Law.

### Mark Schwartz

Mark Schwartz is vice-president and assistant general counsel of Waste Management, Inc. As the leading provider of comprehensive waste management services, Waste Management serves municipal, commercial, industrial, and residential customers throughout North America. Headquartered in Houston, the company's network of operations includes 284 active landfill disposal sites, 16 waste-to-energy plants, 73 landfill gas-to-energy facilities, 160 recycling plants, 293 transfer stations, and more than 1,400 collection facilities, employing approximately 54,000 employees. Mr. Schwartz is responsible for the company's employment law, labor relations, and benefits law functions.

Prior to his position at Waste Management, Mr. Schwartz was the office managing shareholder of the Houston office of Littler Mendelson, a nationwide law firm specializing in labor law and employment litigation. He was extensively involved in all aspects of employment-related litigation, including matters arising under federal and state laws prohibiting discrimination, claims of wrongful termination arising under both common law and contract, sexual harassment claims, and arbitration of labor disputes. Mr. Schwartz also represented employers before such agencies as the Equal Employment Opportunity Commission, the Texas Commission on Human Rights, the Department of Labor, and the National Labor Relations Board.

Mr. Schwartz is a member of the labor law section of the State Bar of Texas.

Mr. Schwartz graduated with his BS, *magna cum laude*, from Texas A & M University and with his JD, *cum laude*, from Baylor University School of Law.

### Bernard R. Siskin, Ph.D.

*Managing Director*  
LECG, LLC

## A PAINLESS PRIMER ON STATISTICAL EVIDENCE IN EMPLOYMENT DISCRIMINATION CASES

By Allan G. King, Littler Mendelson, P.C.

- I. INTRODUCTION
- II. DATA
  - A. Internal Data Sources
    - 1. Data Regarding "Stocks" That Are Used Inappropriately To Address Questions Regarding "Flows"
    - 2. Data That Are Inaccurate
    - 3. Data That Are Conclusory
    - 4. Data That Are Aggregated Improperly
    - 5. Data That Are Incomplete
  - B. External Data Sources
    - 1. The Boundaries Defining The External Source May Be Inappropriate To The Geography Relevant To The Litigation
    - 2. The Legal Environment That Applies To The Subjects Of The External Survey May Differ From That Relevant To The Employees In Question
    - 3. External Data May Be Too Aggregated
    - 4. The Population Relevant To The Litigation May Not Match The Population Surveyed By The External Data Source
- III. MODEL SPECIFICATION
  - A. Grouping Issues
  - B. The Dependent Variable
  - C. The Independent Variables
  - D. Omitted Variables May Bias Results
  - E. An Over-Inclusive Model May Be Biased
  - F. No Fishing Permitted
- IV. MODEL ESTIMATION
  - A. The "80-Percent Rule" Is Misleading
  - B. Probability-Based Measures Are Better
- V. DRAWING INFERENCES FROM STATISTICS
- VI. CONCLUSION

## I. INTRODUCTION

Statistics are the evidentiary keystone in cases alleging disparate impact or a pattern and practice of discrimination. See, e.g., *Watson v. Fort Worth Bank and Trust Co.*, 487 U.S. 977 (1988), and *International Bhd. of Teamsters v. United States*, 431 U.S. 324 (1977). These legal theories, in turn, are at the heart of successful employment class actions. C.f., *Smith v. Texaco, Inc.*, 263 F.3d 394 (5<sup>th</sup> Cir. 2001).

The four components of a statistical analysis addressing disparities in employment are:

- A. Data.
- B. Model Specification.
- C. Model Estimation.
- D. Drawing Inferences from Estimated Results.

See generally, D.H. Kaye and D.A. Freedman, "Reference Guide on Statistics," in Federal Judicial Center, *Reference Manual on Scientific Evidence* (1994), p. 331 *passim*. This outline acquaints attorneys with the pitfalls in using statistical evidence and the hallmarks of sound methodology.

## II. DATA

The overarching question is whether the data relied upon permit a statistician to address adequately the salient issues in the case. Because data rarely are collected with litigation in mind, attorneys must determine which data to provide to their experts and ensure that these data permit their experts to reach sound conclusions. Data may be classified as either *internal* or *external* to the employer.

### A. Internal Data Sources

These typically consist of the employer's payroll and personnel records. In addition, the employer may maintain records, such as EEO-1 reports and affirmative action plans, to comply with government regulations. See 29 CFR §1602.7 and 41 CFR Part 60.2. But, because these data serve purposes that are not litigation-related, they may not be suitable, unless modified or supplemented, for addressing questions that are material to the litigation. The following are examples of how data can be mismatched to the relevant questions.

#### 1. Data Regarding "Stocks" That Are Used Inappropriately To Address Questions Regarding "Flows"

This is a generic problem that leads to flawed "apples and oranges" comparisons. "Stocks" are quantities that can be measured at a particular moment in time. For example, the "number of female employees" is a quantity that can be defined only at a particular point in time, because this number may change as a result of hiring, promotions, quits, and so on. Accordingly, a "stock" is properly defined in terms of *both* a quantity and a specific date.

"Flows," in contrast, are measured *between* two points in time. For example, "the number of employees hired" requires the time *interval* to be defined within which hires can be counted. Therefore, it is appropriate to say, for instance, that 500 employees were hired between January and June. Flows (e.g., new hires) therefore lead to *changes* in stocks.

As a result, it generally is inappropriate to compare stocks and flows in order to draw inferences regarding discrimination. For example, a "snapshot" of employment on a particular date, as presented in an EEO-1 report, usually will not shed light on an employer's hiring practices. The reason is that the number and characteristics of persons employed on a particular date, say December 31, 2001, reflects everyone who was employed on December 31, 2000, and

everyone who was hired, terminated (voluntarily or involuntarily), or promoted during the year. Because of these various "flows" into and out of employment, it is not possible, based solely on employment data (a "stock"), to identify a particular "flow" (hiring) as the cause of any disparity. The same conclusion would apply to inferences regarding terminations and promotions. See, e.g., *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989).

## 2. Data That Are Inaccurate

Not all data retained by employers are relied upon regularly for business purposes. Therefore, the accuracy of these data may not be checked regularly and inaccuracies may be of little consequence to the business. For example, payroll data typically are relied upon to pay employees. If there are errors, these are likely to be caught and corrected either by the employee or management. In contrast, data regarding the previous positions held by an employee, or the employee's gender, may not be used for any purpose. As a result, errors that go undetected can persist in these databases and prove embarrassing at trial, particularly if they relate to the plaintiff or plaintiffs.

## 3. Data That Are Conclusory

Some databases permit employers to characterize an employee's status in ways that may not correspond to the relevant legal definition. For example, an employer's database may characterize a job change as a "promotion" although there was no change in responsibilities or pay. Alternatively, employees may be characterized as "employed," in order to continue their eligibility for benefits, when in fact, they no longer report for work and they have been informed that their job has been eliminated. See *Thurman v. Sears, Roebuck & Co.*, 952 F.2d 128 (5<sup>th</sup> Cir), *cert. denied*, 506 U.S. 845 (1992). Most obviously, employees may be classified as "exempt" from the FLSA's overtime provisions when they are not.

## 4. Data That Are Aggregated Improperly

Because employers compile data with purposes other than litigation in mind, the workforce may be grouped into categories that are either too aggregated or too detailed for meaningful study. For example, EEO-1 reports group employees into broad occupational classes such as "officials and managers" and "professionals." But the former group might include the company's CEO as well as the night-shift manager. "Professionals" includes lawyers and physicians. Consequently, these classifications combine employees who are unlikely to be judged "similarly situated," or compared appropriately.

At the opposite extreme are job classifications that are excessively detailed. Among some employers job titles have proliferated to such an extent that virtually each employee has a unique title. As a result, it may not be possible to analyze employment patterns meaningfully using the job classifications as they exist in the employer's database.

## 5. Data That Are Incomplete

Omitted or incomplete data may be critical to litigation. One example may be an employee's educational attainment. This may be information an employer will rarely consult when making personnel decisions, thus if it is missing it easily may go unnoticed. Yet, "education" may provide a powerful explanation of why one employee failed an employer's aptitude test and another passed easily.

Although there are statistical techniques for coping with missing data, it may be preferable to supplement the database with additional information regarding incomplete records. See generally D. Baldus and J. Cole, *Statistical Proof of Discrimination* (1980), Chs. 4 and 5. In any

event, this is another dimension of the employer's data that should not escape an attorney's attention.

A database also may be incomplete because it omits entirely certain *categories* of important information. For example, an employer may compensate its employees by paying both a salary and a bonus. However, these components of compensation need not be recorded in the same database. Particularly, if the "bonus" is part of a stock plan or other benefit package, it may be omitted from a "payroll" database that primarily is used to generate weekly or monthly paychecks. Obviously, an analysis that is predicated solely on the payroll database would miss an important component of inequality.

## **B. External Data Sources**

The primary characteristic of external data is that, with few exceptions, they do not relate exclusively to the defendant or its workforce. Although proxy data sources may provide highly accurate descriptions of the population or entity that is its subject, the proxy population may differ markedly from the plaintiffs or defendants in the particular lawsuit. Additionally, because the purpose of the external survey is likely to differ from how it is used in litigation, it is important to know whether its quirks or definitions are compatible with its use in litigation.

### 1. The Boundaries Defining The External Source May Be Inappropriate To The Geography Relevant To The Litigation

For example, the Census reports employment in metropolitan areas throughout the nation. The employer, in question, although located within the metropolitan area, may be situated where it draws a significant share of its labor force from an adjacent rural county that is omitted from the Census definition of the metropolitan area. Although the Census may be exceedingly accurate, this measure would mischaracterize the labor market relevant to this employer. See, e.g. *EEOC v. Olson's Dairy Queens, Inc.*, 989 F.2d 165 (5<sup>th</sup> Cir. 1993).

### 2. The Legal Environment That Applies To The Subjects Of The External Survey May Differ From That Relevant To The Employees In Question

For example, the U.S. Department of Labor regularly surveys business establishments and tracks trends in the rate of growth in earnings. These data, however, include payments for overtime work, which would not apply to a group of plaintiffs in "exempt" positions, and therefore may not be the best source of information regarding trends in their salaries.

### 3. External Data May Be Too Aggregated

The detail provided in most publicly-available data reflects privacy concerns that limit the permissible extent of disaggregation. That is, if the Census were to publish the income of the doctor in a small town with only one physician, it necessarily would be disclosing the income of that particular physician. To avoid this, the physician is likely to be included among the larger group of "professionals," which is helpful in protecting the physician's privacy but not in meaningfully defining the labor market. Consequently, the level of detail provided by external data may be too gross for litigation purposes.

An example is provided by *EEOC v. The Mansion, Turtle Creek*, 70 FAIR EMP. PRAC. CASES 899 (N.D. Tex. 1995). The EEOC alleged that this luxury hotel was discriminating against females who sought to become waitpersons in the hotel's dining room. As support, the EEOC relied upon Census data reporting the gender composition of waitpersons in the Dallas metropolitan area. The district court rejected this evidence, finding that the Census job

classification was overly broad because it did not distinguish those waitpersons employed in "fine dining" facilities.

#### 4. The Population Relevant To The Litigation May Not Match The Population Surveyed By The External Data Source

Courts generally prefer analyses of hiring patterns to be based on applicant flow data, *Anderson v. Douglas & Lomason Co.*, 26 F.3d 1277, 1287 (5<sup>th</sup> Cir. 1994), but sometimes there is no choice but to rely on external data. However, even when job classifications in the external data are finely distinguished, there still may be instances in which comparisons are misleading.

Suppose the question is whether a particular firm discriminates in hiring female attorneys. One might be tempted to compare the fraction of female attorneys among those hired to the percentage reported in the appropriate geographic area by some external data source, perhaps the Census or the state bar. This comparison may be misleading if recent cohorts of law school graduates include a higher proportion of females. Because job applicants disproportionately reflect more recent cohorts, a comparison based on the entire group of attorneys will understate the female share of applicants.

### III. MODEL SPECIFICATION

A statistical model necessarily abstracts from the details that make up the decision-making process at issue. A statistician must select from the great many factors that influence whether to hire an applicant, in what position an applicant should be hired, and the rate of pay, those relatively few that are both measurable and material. In addition, a statistical model should reflect precisely *how* those selected characteristics influence decision-making. For example, a qualification either may be a necessary requirement for a particular job or it may simply be a credential that enhances an employee's job prospects. A different statistical approach may be required in each case.

Thus, model specification encompasses a variety of considerations. These include defining the ultimate employment action (the dependent variable), such as hiring, pay, termination, and so on, distinguishing those factors material to the decision from those deemed immaterial, grouping data to define similarly-situated groups of employees, and specifying how the various factors that have been identified interact with each other (*e.g.*, are they additive or multiplicative).

#### A. Grouping Issues

The facts of the case also will affect how the data that are included in the statistical model are grouped. For example, if the case involves layoffs, and the facts establish that layoff decisions are made independently at each of a company's various facilities, it would be a mis-specification to group together all employees, from all facilities, as if they were equally at risk of selection. Instead, it is necessary to define distinct pools of employees at each facility.

For example, suppose 60 percent of the employees at Plant A are age 40 or over, but only 30 percent at Plant B. If equal numbers are employed at each plant, then, in total, 45 percent of all employees are in the protected age group. However, a layoff that disproportionately affects Plant A, for reasons that are independent of age, may result in a layoff that includes a disproportionate number of older workers, even if older workers are selected in proportion to their presence in each plant. This disparity results solely because layoffs were made predominately from Plant A.

A similar issue arises with respect to selections in the same locations. Just as a baseball team requires a player at each of nine positions, most employers, even those faced with layoffs,

require a minimum level of staffing in a variety of positions to function effectively. As a result, it is rarely the case that all employees are equally likely candidates for layoff irrespective of the jobs they hold. Rather, employees in certain positions may be indispensable to future operations, whereas employees in other jobs are less essential. Once again, it would be improper to group together employees in these various positions, as if they were equally likely candidates for layoff.

### **B. The Dependent Variable**

The dependent variable typically is the ultimate employment action that is to be "explained" by the statistical analysis. It usually will be either an indicator variable (yes or no), such as whether a given employee was hired, or a quantity, such as an employee's rate of pay. The dependent variable is in some respects the most important aspect of model specification because unless this variable relates to a material issue in the case – something that relates directly to liability, the statistical analysis may be irrelevant.

Consider an example in which a plaintiff attacks the allegedly discriminatory nature of a particular test. The specific allegation may be that because this test is biased, disproportionately few members of some protected group are hired. There are at least two ways to specify a model that is consistent with this claim. One is to define the dependent variable as whether or not an applicant passed the test. A second approach is to define the dependent variable as whether or not the applicant was hired.

Although the complaint alleges that failing the test results in a failure to obtain the job, it does not follow that all who were not hired necessarily failed the test. Accordingly, a model that specifies the *hiring* decision as a dependent variable may not be probative of whether the test itself is discriminatory. As a result, a court may disregard this statistical analysis as evidence of the discriminatory nature of the test. *Cf. Connecticut v. Teal*, 457 U.S. 440 (1982). Thus, the erroneous specification of the dependent variable may doom the entire statistical analysis.

A complementary question is *how* the dependent variable should be measured. For example, suppose the allegation is that there has been discrimination in selections for layoffs. Although it may seem straightforward that the dependent variable ought to be whether or not a given employee was laid off, the way that question is answered depends on the timeframe that applies. Consider two alternatives. In the first, the question is whether a particular employee was laid off at any time during a given year, whereas in the second the question is whether the employee was laid off during a specified period of time, say a particular week. Whether an employee is designated as "laid off" likely will differ with each choice. Unless layoffs throughout the year were done in very much the same way, the results will depend upon how the dependent variable is defined. The better specification will depend, of course, on the issue that is most relevant to the litigation.

### **C. The Independent Variables**

The independent variables are those factors that are thought to materially influence the dependent variable or the ultimate employment action at issue. This must include some indicator of the protected category (*e.g.*, age, gender, and so on), to test whether membership in a protected group is associated with a significant disparity in the employment action measured by the dependent variable. The other variables to be considered should be dictated by the facts of the case. Thus, what factors should be considered as independent variables is a fact question, and not purely a statistical matter.



For example, suppose one is studying differences in compensation. If the compensation system at issue purports to be a piece-rate system, then it would be appropriate to include as an independent variable a measure of the production attributable to each employee. On the other hand, if employment is governed by a collective bargaining agreement that bases an employee's rate of pay on seniority, then "seniority" and not "production" ought to be included as an independent variable.

Most jobs are less structured than these hypotheticals. As a result, it may be less clear precisely which variables should be included in the statistical model and which may be omitted without biasing the results. Continuing with the example of compensation based upon the seniority, it may also be relevant how long that employee has been in his present job classification. But what if the company's database does not include that information? Can an informative and unbiased statistical analysis still be performed?

#### **D. Omitted Variables May Bias Results**

This raises the issue of "omitted variables." It is rarely the case that information regarding all of the non-discriminatory factors for which one would like to control is at hand. Does the absence of this information destroy the value of any statistical analysis? The answer depends upon whether the variable that is omitted is materially related to the dependent variable, and most importantly, whether the omitted variable is correlated with membership in the protected category.

Suppose that for the past few years a company has made a concerted effort to advance the careers of female employees. As a consequence, many women have been newly-promoted into better paying positions. Suppose it also the case that this company rewards individuals, in part, depending upon how long they have occupied their present positions. Accordingly, women are more likely to be recently promoted, and as a result of their short tenure in these positions they will generally be paid less than males in these same jobs.

If data regarding an employee's "time in job" is omitted from the statistical model, the model will mistakenly attribute to an employee's gender effects that properly should be attributed to the employee's job tenure. Thus, because women on average will have shorter tenure, and shorter tenure typically will result in lower pay, the model will report a correlation between gender and lower pay that properly should be attributed to tenure. Although a statistical model need not include all possible controls, a bias will be introduced if it omits factors that are correlated with both the dependent variable and an employee's protected group status. See, e.g., *Garza v. City of Los Angeles*, 756 F.Supp. 1298, 1234 (C.D. Cal. 1991).

#### **E. An Over-Inclusive Model May Be Biased**

A statistical model may be over-inclusive and erroneously include as an independent variable a factor that should properly be omitted. For example, suppose we again consider the case in which gender discrimination in pay is alleged. Suppose the statistical model now properly includes a variable that measures job tenure, but includes in addition an employee's rating on a performance evaluation. Further, when performance evaluations are included, no gender disparity is present. Is this an appropriate specification of the model?

If a plaintiff is contending that the alleged pay disparity results from a pervasive, sexist bias within the company, then arguably that same bias may taint performance evaluations. By this reasoning the set of independent variables is over-inclusive, because it contains a variable, performance, that reflects the same gender bias that exists with respect to pay. Therefore, this theory suggests that performance should be excluded as an explanatory factor.

## F. No Fishing Permitted

The rules of statistical inference, to be discussed later, assume the existence of a statistical model that is based upon the information or facts that have been adduced. In other words, if a promotional decision exclusively depends upon a candidate's score on a particular test, it would be inappropriate to include as an independent variable information regarding that candidate's grade-point average. On the other hand, if the selection process also reflects the score the candidate receives in an interview, it would be a specification error to exclude that information from the statistical model. Because disputes about the underlying facts translate to arguments about the proper specification of the statistical model, a "fact issue" may exist about the proper specification of the model. Because "the facts" may be malleable, it may be tempting to "reverse-engineer" the statistical model, letting it identify the "facts" that maximize or minimize any disparities.

This result-oriented approach is impermissible and may lead to portraying the statistical evidence in a false light. The problem occurs because standard statistical tests assume the model is specified independently of the data that is used to estimate the model. Consequently, if a statistician purports to assess the significance of the results claimed for a particular model, without disclosing the variety of alternative models that were estimated in the course of arriving at the specified model, then the statistical properties claimed for that model will be misleading.

Consider an example. It is rare for a fair coin to be flipped 10 times and land each time on "heads." In fact, the probability is  $.5^{10}$  or .00001. On the other hand, if a coin is flipped 10,000 times, the probability that somewhere in the process there will be a run of 10 consecutive heads is far greater. Thus, were a statistician to report the rare occurrence of 10 consecutive "heads," without disclosing this was obtained in the course of 10,000 coin flips, he or she would have seriously misrepresented the significance of the occurrence.

An obvious opportunity for such experimentation is provided by age discrimination cases. Although the statute prohibits discrimination against those 40 and over, some courts recognize that discrimination claims also may be brought by subgroups of the protected category, (e.g., employees 50 years of age and older). *Rose v. Wells Fargo & Co.*, 902 F.2d 1470 (9th Cir. 1990). As a consequence, a plaintiff may be tempted to experiment with a variety of potential subgroups, conceivably every year of age over 40. A statistician could perform an analysis for those 40 years of age and older, 41 years of age and older, 42 years of age and older, and so on. Then, having ascertained the particular subgroup that maximizes any disparity between the subgroup and other employees, the plaintiff may cast himself as the victim of discrimination against that subgroup. Then, using the same data by which that subgroup was identified, the statistician can attempt to establish the merits the claim itself.

That approach, known as "data mining," is illegitimate. The statistical tests reported in standard software packages indicate the probability of finding a disparity against a *pre-specified* group or subgroup (e.g., persons 40 years of age and older). However, the probability of finding a statistically significant disparity against *some* subgroup is vastly greater. Thus, if a party presents only the results regarding the subgroup that best fits its case, and fails to report the insignificant or less significant results obtained for all the other subgroups on whom he experimented, the reported results will be seriously misleading. Accordingly, data mining can succeed as a strategy only if it is undisclosed.

## IV. MODEL ESTIMATION

There are typically three fundamental parameters that must be estimated in employment discrimination cases. These are (1) the magnitude of any alleged disparity, (2) the degree to

which that disparity is within limits consistent with a hypothesis of randomness, and (3) how well the statistical model characterizes the underlying data.

The measured disparity between the protected group and others typically depends upon the particular specification of the model. In the crudest of models, one simply calculates, for example, the difference in pay between two groups, not controlling for any influences apart from the protected characteristic. These crude comparisons often generate the largest disparity because they fail to consider factors, apart from a particular demographic characteristic, that also are associated with differences between employees. This is the "omitted variable" problem discussed above.

Once the magnitude of any disparity is determined, the next question is whether the disparity is large enough to be meaningful or "substantial." Even a random selection process will not produce proportional results each and every time. Going back to our example of coin flipping, a fair coin will not produce heads or tails precisely fifty percent of the time, in each and every set of trials. Therefore, it is relevant to ask whether the disparity that is observed is larger than what reasonably would occur if differences solely were attributable to chance.

#### **A. The "80-Percent Rule" Is Misleading**

Although it may seem at first blush that this "greater than chance" standard could be drawn with bright-line precision, that is not the case. Indeed, that precisely is the problem with relying on the "80-percent rule" that is associated with the EEOC's regulations. 29 CFR § 1607.4 (D). The difficulty with this formulation is that the whether any disparity is consistent with chance depends, at least in part, on the number of occurrences that are observed. For example, suppose a coin is flipped just twice and two heads result. According to the "80-percent rule," this departure should be probative of an unfair coin because "heads" occurred twice instead of just once. Yet, the likelihood of obtaining two heads in two flips is 25 percent, not at all a rare occurrence.

More generally, whether any given deviation from proportionality is consistent with chance, will depend, in part, on the number of events (*e.g.*, hires, layoffs, and so on). As a result, an absolute standard, such as the "80-percent rule", will falsely identify deviations as material when samples are small, and understate the probative value of deviations when samples are large. As a consequence, many courts have disparaged this rule. *See, e.g., Clady v. Los Angeles Co.*, 770 F.2d 1421, 1428 (9<sup>th</sup> Cir. 1985), *cert. denied*, 475 U.S. 1109 (1986); and *NAACP v. City of Mansfield*, 866 F.2d 162, 168 (6<sup>th</sup> Cir. 1989).

#### **B. Probability-Based Measures Are Better**

Instead, courts prefer relative measures that reflect the *probability* that any observed disparity could have occurred as a result of a specified random process. A convenient point of reference is the bell-shaped curve of the normal distribution of probabilities. The parameters of this curve are its mean and standard deviation. Once these are known, the probability of any disparity occurring can be ascertained as well. Courts have generally required proof that an outcome would occur under a random process no more than five percent of the time, before finding that that a disparity is probative of discrimination. If outcomes are randomly distributed, this corresponds to disparities that are at least two standard deviations (actually 1.96) from the mean. *Hazelwood Sch. Dist. v. United States*, 443 U.S. 299 (1977).

Although there is a one-to-one relationship between the number of standard deviations and the probability of the occurrence under the specified random process if the distribution of outcomes is normally distributed, the same relationship does not hold for every distribution. Particularly, if a sample is small, then the normal distribution may *not* approximate the

distribution of outcomes. In that event, mechanically applying rules regarding standard deviations may produce seriously misleading results. Therefore, Courts often express skepticism about the value of statistical disparities that are obtained from small samples of data. See, e.g., *Turner v. Texas Instruments, Inc.*, 555 F.2d 1251, 1257 (5<sup>th</sup> Cir. 1977).

Although that skepticism is well-founded if directed at tests that require normality, it is too sweeping when applied to statistical procedures that do not assume normality. For example, Fisher's exact test does not depend on an assumption of normality and, as the name suggests, it permits a statistician to exactly calculate the probability of observed disparities. *Sheehan v. Daily Racing Form, Inc.*, 104 F.3d 940 (7<sup>th</sup> Cir.), cert. denied, 521 U.S. 1104 (1997). Although small samples make it less likely that disparities of any given magnitude will be judged significant, if a significant disparity is obtained for a small sample, it is no less probative of discrimination merely because the sample is small.

The third facet of the estimating procedure is an overall assessment of how well the estimated statistical model fits the data. If the results predicted by the model correspond closely to the outcomes that actually are observed, a statistician can be more confident that the model captures those variables relevant to the employer's decision making process. Alternatively, if there is substantial disagreement between the results predicted by the model and the actual outcomes, this signals that relevant factors may be omitted, or else that the model misstates how the included variables interact with each other. In either case, a poorly-fitting model raises questions about the completeness, and perhaps the accuracy, of the results obtained.

## V. DRAWING INFERENCES FROM STATISTICS

It is a truism that statistics can never "prove" the existence of discrimination. Rather, at most, statistics can lead a fact-finder to reject "randomness" as an explanation of any disparity. This leaves discrimination, along with other non-discriminatory factors not explicitly taken into account by the statistical model, as potential explanations.

The probability that any observed disparity could have occurred under a specified random process is known as its "p-value." Unlike the 80-percent rule, and the "2 or 3 standard deviation" criterion of *Hazelwood*, the p-value provides a standard that is unaffected by the size of a sample or the particular shape of the distribution of outcomes.

It is conventional for courts to insist that proffered statistics have a p-value of no more than five percent to be considered probative evidence. See, e.g., *W.G. Bennett v. Total Minatome Corp.*, 138 F.3d 1053 (5<sup>th</sup> Cir. 1998). Recently, however, Judge Posner has suggested that this .05 level of significance is arbitrary, being a carryover from the conventions applied by academic journals to limit the volume of scholarly papers that are published. Judge Posner suggests that statistics associated with p-values greater than .05 therefore might be admissible, with evidence of the p-value affecting the weight as opposed to the admissibility of those statistics. *Kadas v. MCI Systemhouse Corp.*, 255 F.3d 359, 362 (7<sup>th</sup> Cir. 2001). Judge Posner's comments reflect the trend away from the mechanical application of statistical criteria in favor of a more thoughtful, case-by-case assessment of statistical evidence.

The problem caused by "data mining" or "fishing" also can be understood in terms of p-values. The *Hazelwood* rule regarding "2 or 3 standard deviations" reflects the fact that if outcomes are normally distributed, those that lie two or more standard deviations from the mean, or the expected value, occur with a frequency no greater than five percent. However, this and similar criteria are valid only for a given, pre-specified statistical model. If, on the other hand, the reported statistical results are merely the end-product of a search among a variety of statistical models, the results are misleading. In such cases, the appropriate p-value is the

probability of finding *any* significant disparity in several statistical trials. This is vastly different than the p-value calculated for a given statistical trial, which provides the usual *Hazelwood* criterion. Accordingly, p-values obtained through an elaborate search process for significant results should be deeply discounted, and accorded far less weight.

The maxim “garbage in, garbage out” is a useful reminder that if the statistical results are premised upon “apples and oranges” comparisons or result from numerous “experiments” with the data, then the resulting p-values may be largely irrelevant.

Additionally, some courts have circumscribed the use of statistical evidence. For example, if a plaintiff alleges a theory of disparate treatment, some courts have held that while statistical evidence may be useful in establishing a *prima facie* case, once the defendant has articulated a legitimate, non-discriminatory reason for its actions, a plaintiff no longer may rely on that same statistical evidence to prove the employer’s explanation is pretextual. *Barnes v. GenCORP Inc.*, 896 F.2d 1457 (6<sup>th</sup> Cir.), *cert. denied*, 498 U.S. 878 (1990). Thus, if a plaintiff alleges that he was denied a promotion because of his membership in a protected group, and supports his claim by adducing evidence that disproportionately few members of his group received promotions, the employer may respond with evidence that the particular plaintiff never applied for the promotion. *Barnes* suggests that this explanation cannot be rebutted merely by reiterating that there is a deficit in the promotions of that protected group. However, that holding does not stand for the proposition that statistics can never rebut an employer’s explanation. In this hypothetical, a plaintiff might use statistics to demonstrate that a significant fraction of those promoted also failed to apply for the position, thereby negating the employer’s explanation.

## VI. CONCLUSION

I have attempted to distill the principal lessons of this outline into Ten Commandments of Statistical Analysis. They are as follows:

1. Verify that the data accurately reflect the issues at hand.
2. Make sure your expert understands the facts and salient issues in the case.
3. Never compare apples and oranges.
4. Analyze flows not stocks.
5. Do not data mine.
6. Never rely on the 80-percent rule.
7. Do not discount the value of small samples.
8. Rely on p-values, not standard deviations.
9. Do not “reverse-engineer” factual allegations.
10. Do not mistake statistical significance for “substantial disparities.”

## THE THEORY OF DISPARATE IMPACT

**By Allan G. King, Littler Mendelson, P.C.**

Targets practices that are fair in form but discriminatory in their impact.

### **I. HISTORICAL ORIGINS.**

- A. *Griggs v. Duke Power Co.*, 401 U.S. 424, 430 (1971) (“practices, procedures, or tests neutral on their face, and even neutral in terms of intent, cannot be maintained if they operate to ‘freeze’ the status quo of prior discriminatory practices”).
- B. *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 422 (1975) (“Title VII is not concerned with the employer’s ‘good intent or absence of discriminatory intent’”); “[E]ven validated tests might be a ‘pretext’ for discrimination in light of alternative selection procedures available to the Company.” *Id.* at 436.
- C. *Connecticut v. Teal*, 457 U.S. 440 (1982) (disparate impact cannot be eradicated by parity at the “bottom line”).
- D. *Watson v. Fort Worth National Bank & Trust Co.*, 487 U.S. 977 (1988) (disparate impact model may be applied to subjective decision making criteria); Plaintiff is “responsible for isolating and identifying the specific employment practices that are allegedly responsible for any observed statistical disparities.”
- E. *Wards Cove Packing v. Atonio*, 490 U.S. 642 (1989) (burden of persuasion remains with plaintiff; plaintiff must isolate and identify the particular employment practices causing the disparate impact).

### **II. CIVIL RIGHTS ACT OF 1991**, codified in 42 U.S.C. §§ 1981 *et seq.*

- A. Codified disparate impact under Title VII and ADA.
- B. Excluded disparate impact from provisions regarding jury trials and compensatory or punitive damages.
- C. Reallocated burdens of proof relative to *Wards Cove*.
- D. Restated employer’s burden in response to prima facie case.
- E. Permits employer’s practices to be cumulated in cases where plaintiff can demonstrate that the elements of an employer’s decision-making process are incapable of separation.

### **III. THE PRIMA FACIE CASE.**

Plaintiff must:

- (1) identify a particular employment practice.
- (2) that causes.
- (3) a substantial disparate impact.
- (4) on a class in an employer’s workforce.

*Pouncy v. Prudential Ins. Co. of Am.*, 668 F.2d 795, 800 (5th Cir. 1982); *Moore v. Southwestern Bell Tel. Co.*, 593 F.2d 607, 608 (5th Cir. 1979) (selection device must have a “substantially disproportionate impact”) (quoting *Washington v. Davis*, 426 U.S. 229, 247 (1976)).

#### A. Identify with Particularity

- 1. “a complaining party [must] demonstrate that a respondent uses a particular employment practice . . .” 42 U.S.C. § 2000e - 2(k)(1)(A)(i).
- 2. “the complaining party shall demonstrate that each particular challenged employment practice causes a disparate impact, except if the complaining party

can demonstrate to the court that the elements of a respondent's decision making process are not capable of separation for analysis, the decision making process may be analyzed as one employment practice." *Id.* § 2000e-2(k)(1)(B)(i).

3. "When a decision making process includes particular functionally-integrated practices which are components of the same criterion, standard, method of administration or test, [they] may be analyzed as one employment practice." Interpretative Memorandum, 137 Cong. Rec. 15,276 (Oct. 25, 1991) (exclusive legislative history).
4. "The discriminatory impact model of proof in an employment discrimination case is not, however, the appropriate vehicle from which to launch a wide ranging attack on the cumulative effect of a company's employment practices . . . . The disparate impact model applies only when an employer has instituted a specific procedure, usually a selection criterion for employment, that can be shown to have a causal connection to a class based imbalance in the work force." *Pouncy*, 668 F.2d at 800.
5. "disparate impact analysis is in principle no less applicable to subjective employment criteria than to objective or standardized tests." *Watson*, 487 U.S. at 991.

B. Causation

1. "[T]he plaintiff must offer statistical evidence of a kind and degree sufficient to show that the practice in question has caused the exclusion of applicants for jobs or promotions because of their membership in a protected group." *Id.* at 994.
2. must show "that independent of other factors the employment practices he challenges have caused the . . . imbalance. [An imbalance] might result from any number of causes. Absent proof that the impact is caused by one or more of the challenged employment practices, we do not require the employer to justify the legitimacy of any (or all) employment practices." *Pouncy*, 668 F.2d at 801-02.

C. A Substantial Disparate Impact

1. "Statistics are the "inevitable focus" of disparate impact cases. *Watson*, 487 U.S. at 992.
2. "statistical disparities must be sufficiently substantial that they raise such an inference of causation." *Id.* at 995.
3. The "four-fifths" rule adopted in the EEOC's *Uniform Guidelines on Employee Selection Procedures*, 29 C.F.R. pt. 1607 (1987), "has not provided more than a rule of thumb for the courts." *Watson*, 487 U.S. at 995 n. 3. "But, . . . , they do constitute '[t]he administrative interpretation of the Act by the enforcing agency,' and consequently they are 'entitled to great deference.'" *Albemarle Paper Co.*, 422 U.S. at 431.
4. "we have not suggested that any particular number of 'standard deviations' can determine whether a plaintiff has made out a prima facie case." *Id.*
5. Statistical comparisons must be drawn between members of the appropriately-defined, qualified labor pool (e.g., the population at large is not usually the appropriate pool to whom an employer's workforce should be compared). See *New York Transit Authority v. Beazer*, 440 U.S. 568, 584-87 (1979); *Johnson v. Uncle Ben's, Inc.*, 965 F.2d 1363, 1369 (5th Cir. 1992).

D. On A Class In An Employer's Workforce

It is well-settled that a disparate impact claim can be stated on behalf of groups protected by Title VII and/or the ADA, 42 U.S.C. § 2000e-2(k)(1)(A)(i) and § 12112(b)(6). Circuits are divided as to whether this theory is cognizable under the ADEA, 29 U.S.C. § 621 *et seq.*

1. “By contrast, *Hazen Paper*, indicates that disparate impact theory is not available under ADEA.” *Rhodes v. Guiberson Oil Tools*, 75 F.3d 989, 1004 (5th Cir. 1996) (en banc) (JJ. DeMoss, J. Smith, and Hawkins, concurring in part and dissenting in part).
2. “We have never decided whether a disparate impact theory of liability is available under the ADEA.” *Hazen Paper Co. v. Biggins*, 507 U.S. 604, 113 S.Ct. 1701, 1706 (1993).
3. See also *Ellis v. United Airlines, Inc.*, 73 F.3d 999, 1003 (10th Cir. 1996) (no disparate impact under the ADEA); *EEOC v. Francis W. Parker School*, 41 F.3d 1073 (7th Cir. 1994) (same).
4. *Adams v. Florida Power Corp.*, 255 F.3d 1322 (11th Cir. 2001) (“disparate impact not cognizable under the ADEA”).
5. But see *Criley v. Delta Airlines, Inc.*, 119 F.3d 102, 105 (2d Cir. 1997) (“recognizing disparate impact claim under the ADEA”); *Frank v. United Airlines, Inc.*, 261 F.3d 845, 856 (9th Cir. 2000) (same).

#### IV. **THE EMPLOYER’S BURDEN TO PROVE “BUSINESS NECESSITY.”**

- A. “Business necessity” has been a shifting standard, both in terms of what must be proved and who must prove it.
  1. The challenged practice “must have a manifest relationship to the employment in question.” *Griggs v. Duke Power Co.*, 401 U.S. at 432.
  2. Employer must prove test is “job related.” *Albemarle Paper Co. v. Moody*, 422 U.S. at 425.
  3. The practice was required by the employer’s “legitimate employment goals of safety and efficiency.” *New York City Transit Authority v. Beazer*, 440 U.S. at 587 n. 31.
  4. Employer must meet “its burden of producing evidence that its employment practices are based on legitimate business reasons, . . .” *Watson v. Fort Worth Bank & Trust*, 487 U.S. at 998.
  5. “[T]he dispositive issue is whether a challenged practice serves, in a significant way, the legitimate employment goals of the employer.” *Wards Cove Packing Co., Inc. v. Atonio*, 109 S.Ct. at 2125. “[T]here is no requirement that the challenged practice be “essential” or “indispensable” to the employer’s business to pass muster. . . .” *Id.* at 2126. “[T]he employer carries the burden of producing evidence of a business justification for his employment practice. The burden of persuasion, however, remains with the disparate impact plaintiff.” *Id.*
  6. “[T]he respondent fails to demonstrate that the challenged practice is job related for the position in question and consistent with business necessity.” 42 U.S.C. § 2000e-2(k)(1)(A)(i).
    - (1) The terms “business necessity” and “job related” are intended to reflect the concepts enunciated by the Supreme Court in *Griggs v. Duke Power Co.* . . . , and in other Supreme Court decisions prior to *Wards Cove Packing Co. v. Atonio* . . . Interpretive Mem., 137 Cong. Rec. S 15,474.
    - (2) Draft of the legislation vetoed by President Bush provided that a challenged practice, with respect to employment selection, bear “a significant relationship to the successful performance of the job,” and with respect to other employment decisions, bear “a significant relationship to a manifest business objective of the employer.”
    - (3) “Job related for the position in question” likely narrows the employment goals that will suffice to justify challenged practices and excludes



- justifications such as customer preference, morale, corporate image and convenience, under the *Griggs* standard. 137 Cong. Rec. 9,528.
- (4) Although formal validation studies may suffice where the challenged practice is a test, validation studies are not necessary. *Watson v. Fort Worth Bank & Trust*, 487 U.S. at 998 (employers are not required to introduce formal “validation studies” showing that particular criteria predict actual on-the-job performance). *But see Albemarle Paper Co. v. Moody*, 422 U.S. at 431 (“discriminatory tests are impermissible unless shown by professionally accepted methods, to be ‘predictive or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated’”); cf. *Clady v. County of Los Angeles*, 770 F.2d 1421, 1428-29 (9th Cir. 1985), *cert. denied*, 475 U.S. 1109 (1986) (Uniform Guidelines not legally binding).

**V. PLAINTIFF MAY REBUT BY PROOF OF AN ALTERNATIVE, LESS DISCRIMINATORY ALTERNATIVE**

- A. “[I]t remains open to the complaining party to show that other tests or selection devices, without a similarly undesirable racial effect, would also serve the employer’s legitimate interest. . . . Such a showing would be evidence that the employer was using its tests merely as a ‘pretext’ for discrimination.” *Id.* at 425.
- B. The complaining party makes the demonstration with respect to an alternative employment practice and the respondent refuses to adopt such alternative employment practice. The demonstration of an “alternative employment practice” shall be in accordance with the law as it existed prior to *Ward’s Cove*, 42 U.S.C. §§ 2000e-2(k)(1)(A)(ii) and (2)(k)(1)(C). *Ward’s Cove* suggests “the judiciary should proceed with care before mandating that an employer must adopt a plaintiff’s alternate selection or hiring practice.” 109 S.Ct. at 2127. *But see Watson*, 487 U.S. at 998. (“Factors such as the cost or other burdens of proposed alternative selection devices are relevant in determining whether they would be equally as effective”). *See also Fitzpatrick v. City of Atlanta*, 2 F.3d 1112, 1122 (11th Cir. 1993) (rejecting plaintiffs’ proposed alternative).
- C. There are several unanswered questions regarding what constitutes sufficient proof of an alternative employment practice. For example, how much less discriminatory must the alternative be? What if the alternative is much less discriminatory but slightly less predictive of job performance?

**VI. EXAMPLES OF HOW DISPARATE IMPACT THEORY HAS BEEN APPLIED.**

- A. *Stender v. Lucky Stores, Inc.*, 803 F. Supp. 259 (N.D. Cal. 1992).
1. Allegations  
Race and sex discrimination in initial placement, allocation of overtime, and promotions.
  2. Particular practice challenged.
    - a. ambiguous and subjective decision making practices.
    - b. “Where the system of promotion is pervaded by a lack of uniform criteria, . . . the court is not required to ‘pinpoint particular aspects of the system’ . . .” *Id.* at 335.
    - c. subjective and ambiguous decision making processes are not separable.
  3. Business necessity.
    - a. Defendant’s failed to follow its own policies, and these departures have had a disparate impact.
    - b. “Defendant’s failure to follow its own policies is excepted by neither the business necessity nor the business justification defenses.

4. At least seven expert witnesses testified.
- B. *Graffam v. Scott Paper Co.*, 870 F. Supp. 389 (D. Me. 1994), aff'd, 60 F.3d 809 (1st Cir. 1995).
  1. Allegation.  
Violation of ADEA in reduction of force; no mention of *Hazen Paper*. Court observes that portions of *Wards Cove* were overruled by CRA of 1991, although these amendments were to Title VII, not ADEA.
  2. Challenged practice.
    - a. “the entire subjective decisional practice may be analyzed as one practice.”
    - b. Court permits challenge by group “50 and over,” *Graffam v. Scott Paper Co.*, 848 F. Supp. 1 (D. Me. 1994), *but see Lowe v. Commack Union Free School Dist.*, 886 F.2d 1364, 1362-73 (2d Cir. 1989), *cert. denied*, 494 U.S. 1026 (1990) (subgrouping by age is impermissible).
  3. Job Related and Business Necessity.
    - a. “business necessity” inquires whether the job criteria arise out of a manifest business need--large scale layoffs were due to a business necessity.
    - b. “job related” standard inquires whether there is a correlation between the criteria used and successful job performance. Court found “job relatedness” based upon expert’s testimony that each of the subjective criteria, including interpersonal skills, self-management, performance, and versatility--described job behaviors required in managerial and technical jobs in a manufacturing facility. *Id.* at 402. Court ignored Plaintiff’s criticism that these traits were not actually correlated with the performance of any job. *Id.* at 401.

## VII. PATTERN AND PRACTICE THEORY

Usually focuses on systemic discrimination that allegedly pervades the challenged employment practices.

### A. Burden of Proof

1. Plaintiff must prove by a preponderance of evidence that discrimination is the company’s standard operating procedure—the regular rather than the unusual practice. *International Bhd. of Teamsters*, 431 U.S. 324, 335-36 (1977); *Hazelwood Sch. Dist. v. United States*, 433 U.S. 299 (1977).
2. Proof usually is made by statistical and anecdotal evidence. *Teamsters*, 431 U.S. at 338.
3. If plaintiff makes this showing, the burden shifts to the employer to prove that the adverse action would have occurred in the absence of discriminatory intent. *Franks v. Bowman Transp. Co.*, 424 U.S. 747, 773 (1976).

## MAINTAINING THE CONFIDENTIALITY OF A STATISTICAL AUDIT

*by Mark Schwartz, Vice President and Assistant General Counsel of Waste Management, Inc.*

If the statistical audit is undertaken to avoid litigation, it is unlikely to be deemed in "anticipation of litigation," as the work-product doctrine requires. FED. R. CIV. P. 26(b)(2). Consequently, the attorney-client privilege, and the "self-critical analysis" privilege, where recognized, potentially protect the audit results.

### I. ELEMENTS OF THE ATTORNEY-CLIENT PRIVILEGE

A. The attorney-client privilege applies to communications made between privileged persons in confidence for the purpose of seeking, obtaining, or providing legal assistance to the client. RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS § 68 (2000).

B. The privilege protects communications, not facts. *Cf. Wonneman v. Stratford Securities Co.*, 23 F.R.D. 281, 285 (S.D.N.Y. 1959). Accordingly, it is the attorney's analysis of the data underlying an audit that may be protected, not the data themselves.

C. A communication is not privileged simply because one of the parties is a lawyer. The attorney must be consulted for the purpose of obtaining legal services or advice that a lawyer may perform in his capacity as a lawyer, not in some other capacity. 8 Wright and Miller, *Federal Practice and Procedure*, § 2017, p. 136. Thus, in-house counsel who perform a variety of tasks, some of which may be operational, not legal, should be cautious about directing a statistical audit themselves.

### II. THE CORPORATION AS "CLIENT"

The attorney-client privilege unquestionably applies to corporations. *Upjohn Co. v. United States*, 449 U.S. 383 (1981). The salient question is which corporate communications are deemed to be those of the client.

### III. THE "CONTROL GROUP" TEST

A. An employee's statement is not considered a corporate communication unless the employee is in a position to control or take substantial part in a decision about any actions which the corporation may take upon the advise of an attorney, or if he is an authorized member of a body or group that has that authority. *Harper & Row Publishers, Inc. v. Decker*, 423 F.2d 487 (7th Cir. 1970), *aff'd*, 400 U.S. 348 (1971).

B. The privilege extends to a second tier of employees, below top management, "whose advisory role to top management in a particular area is such that a decision would not normally be made without their advice or opinion, and whose opinion in fact forms the basis of any final decision by those with actual authority." *Id.* at 258. "If such an employee is consulted to determine a legal course of action, that employee's communication is protected from disclosure." *Caremark, Inc. v. Affiliated Computer Services, Inc.*, 192 F.R.D. 263, 265-66 (N.D. Ill. 2000).

C. The "control group" test continues to be applied in Illinois. *Consolidation Coal Co. v.*

*Bucyrus-Erie Co.*, 432 N.E.2d 250 (Ill. 1982) (only an employee whose advisory role to top management...is such that a decision would not normally be made without his advice or opinion is a member of the control group).

D. If the corporation directly retains an outside statistical consultant, the communication will be protected if:

- (1) The non-employee agent served as an advisor to top management of the corporate client;
- (2) The advisory role was such that the corporate principal would not normally have made a decision without the agent's opinion or advice; and
- (3) The agent's opinion or advice in fact formed the basis of the final decision made by those with actual authority within the corporate principal. *Id.* at 267 (approving arrangement where corporation retained an outside accounting firm to assist its outside counsel in providing legal advice to corporate executives).

E. Evidence may be admitted in Illinois that is privileged elsewhere: "Evidence that is privileged under the local law of the state which has the most significant relationship with the communication, but which is not privileged under the local law of the forum, will be admitted unless there is some special reason why the forum policy favoring admission should not be given effect." RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 139(2).

#### IV. THE "SUBJECT MATTER" TEST

A. Emphasizes the attorney's role as a fact-gatherer.

B. Elements of subject-matter test:

- (1) The communication was made for the purpose of securing legal advice;
- (2) The employee making the communication did so at the direction of his corporate supervisor;
- (3) The superior made the request so the corporation could secure legal advice;
- (4) The subject matter of the communication is within the scope of the employee's corporate duties; and
- (5) The communication is not disseminated beyond those with a need to know.

*Diversified Indus., Inc. v. Meredith*, 572 F.2d 596 (8th Cir. 1978) (en banc).

C. The key inquiry remains whether communication was for the purpose of providing legal advice to the corporation. Regarding the statistical audit, the lawyer must play an interpretive role and provide a legal opinion regarding the results and may not be a mere conduit. *See, e.g. United States v. Gurtner*, 474 F.2d 297 (9th Cir. 1973) ("If what is sought is not legal advice but only account service...or if the advice sought is the accountant's rather than the lawyer's, no privilege exists").

## V. THE SELF-CRITICAL ANALYSIS PRIVILEGE

A. Recognized by some courts under circumstances in which confidentiality is essential to the free flow of information...and the free flow information is essential to promote recognized public interests." Note, *The Privilege of Self-Critical Analysis*, 96 Harv. L. Rev. 1083, 1087 (1983).

B. District courts have reached opposite conclusions regarding its applicability to statistical audits of employment practices. See *Barks v. Lockheed-Georgia Co.*, 53 F.R.D. 283 (N.D. Ga. 1971) (protecting evaluation of affirmative action plans); but see *Webb v. Westinghouse Elec. Corp.*, 81 F.R.D. 431 (E.D. Pa. 1978) (requiring material to have been prepared for mandatory government reports, protecting only subjective, evaluative material, not objective data, and balancing these against plaintiff's need for the material); and *Tharp v. Sioyer Steel Corp.*, 149 F.R.D. 177, 181 (S.D. Iowa 1993) (surveying employment discrimination cases and finding 13 courts that rejected the privilege, and nine in which the court recognized the privilege).

C. No circuit court of appeals has recognized the privilege. *Johnson v. United Parcel Service, Inc.*, 206 F.R.D. 686, 689 (M.D. Fla. 2002).

D. The elements of the privilege have been formulated as follows:

- (1) The documents were prepared for mandatory government reports;
- (2) The privilege is being asserted only to protect subjective materials;
- (3) The privilege is not being asserted to protect objective data in the same report;
- (4) Policy favoring exclusion outweighs plaintiff's need for the information.

*Spina v. Forest Preserve of Cook Co.*, 2001 WL 1491524 (N.D. Ill., Nov. 23, 2001).

E. But see the alternative formulation of *Reid v. Lockheed Martin Aeronautics Co.*, 199 F.R.D. 379, 385 (N.D. Ga. 2001).

(1) The information must result from the self-critical analysis of the party seeking protection;

(2) The public must have a strong interest in preserving the free flow of the type of information sought; and

(3) The information must be of the type whose flow would be curtailed if discovery were allowed.

## **VI. STEPS TO ENHANCE CONFIDENTIALITY OF THE STATISTICAL AUDIT OF EMPLOYMENT PRACTICES.**

### **A. Retain Outside Counsel**

This strengthens the evidence that the corporation is seeking legal advice; exclusive reliance on in-house counsel, if that lawyer also has non-legal responsibilities, will weaken claim. *United States Postal Service v. Phelps Dodge Refining Corp.*, 852 F.Supp. 156 (E.D.N.Y. 1994) (“Defining the scope of the privilege for in-house counsel is complicated by the fact that these attorneys frequently have multi-faceted duties that go beyond traditional tasks performed by lawyers”).

### **B. Have Outside Counsel Retain the Statistical Consultant on Behalf of the Client**

The purpose of the statistical consultant is to assist the lawyer in rendering legal advice. The data should be analyzed and interpreted for the benefit of the lawyer, who then analyzes its legal implications for the benefit of the client. *United States v. Kovel*, 296 F.2d. 918 (2d Cir. 1961).

### **C. The Statistical Consultant Should Communicate with Outside Counsel, not the Client**

The attorney may not simply function to procure statistical analysis for the client. “[I]f the advice sought is the accountant’s, no privilege exists.” *Id.*

### **D. Outside Counsel Should Provide Legal Analysis of the Consultant’s Findings, Not Just the Statistics Themselves**

The attorney-client privilege extends to professionals hired to assist the attorney, if:

(1) The professional was consulted, in confidence, for the purpose of obtaining legal advice from the lawyer; and

(2) The communications between the professional, client, and lawyer are reasonably related to the purpose of obtaining confidential legal advice from the lawyer.

*Aull v. Cavalcade Pension Plan*, 185 F.R.D. 618 (D. Colo. 1998) (discussing communications with accountant).

**MANAGING AND USING EMPLOYMENT INFORMATION:  
TODAY'S ARMOR OR TROJAN HORSE?**

**by**

**Bernard R. Siskin, Ph.D.**

**LECG, LLC.**

**MANAGING AND USING EMPLOYMENT INFORMATION:  
TODAY'S ARMOR OR TROJAN HORSE?**

by

**Bernard R. Siskin, Ph.D.**

**LECG, LLC.  
Philadelphia, Pennsylvania**

**I. INTRODUCTION**

Employers face the ever-present dilemma of whether it is better to know or not to know the legal implications of possible evidence of discrimination contained in employment information databases. In the computer age, it is relatively easy for plaintiffs and defendants to analyze the existing computerized information. Thus, the question is, do we or don't we analyze the data we have on hand? The answer is an unqualified yes and no. This paper will discuss what "looking at your data" entails, the goal of such an effort, its potential advantages and the commitment necessary on the part of management to mount such a proactive effort.

To undertake such a proactive effort to identify possible existing discriminatory practices and alter them for the future requires a commitment to go beyond simply looking at the implications of your existing databases. It requires a commitment to: (1) ask hard questions of your decision makers, if necessary, (2) collect new data which is not really needed to run your business, but which may be needed to "defend" your decision processes, and (3) change, monitor and/or validate your decision processes in the future, if necessary. If a company is willing to make such a total commitment, a proactive statistical analysis will act as armor against claims of discrimination and may even improve the employer's employment processes. However, if management is committed only to looking at the implications of its existing data and does not commit to follow up on the results of its prima facie analysis with changes in its



employment practices, if necessary, then such a proactive study will serve as the Trojan horse plaintiffs use to topple your employment practices.

The bases for this opinion are delineated in the rest of this paper.

We will take a somewhat backwards approach by discussing:

- the elements of prima facie statistical evidence;
- the limits of statistical analysis; and
- statistical responses available to a defendant.

With this background, hopefully you will see how statistical analysis may be effectively utilized before the fact rather than after a claim is filed and after an affidavit from plaintiff's expert statistician has been submitted. The message which we hope to deliver is that a thorough analysis of the selection process will:

- force the employer to think about the appropriateness of the criteria utilized for decision-making;
- determine whether such criteria are being uniformly applied and identify potential problem areas;
- lead to a credible "paper trail" for any future defense of the process which may be required;
- demonstrate management's concern for fairness; and
- reduce the opportunity for a plaintiff to allege that unfettered discretion is the common practice, in order to justify a class in a potential class action.

Of necessity, we must discuss some basic statistical concepts and methods. However, for those of you for whom Statistics 101 is a distant (and likely unpleasant) memory, take heart . . . you will not be required to appreciate the subtleties of our trade to recognize the value of early statistical intervention.

## II. DISPARATE TREATMENT AND IMPACT THEORIES OF DISCRIMINATION AND THE PRIMA FACIE STATISTICAL CASE

There are two legal theories of discrimination (impact and treatment) as well as two types of actions (individual and class). The role and use of statistical evidence varies significantly based on the theory advanced and the type of action involved.

Disparate impact claims are, by definition, class actions, normally involving protected groups defined on the basis of race or gender. Such claims concern a challenge to a policy or practice which, while applied neutrally, has the unintended effect of disproportionately harming or disadvantaging a protected group. Discriminatory intent is not at issue in such claims. As such, the basic question is purely statistical: did the neutrally applied employment practice have a disparate impact upon the protected group? The Civil Rights Act of 1991 confirmed the Court's view of what constitutes a prima facie disparate impact (outlined in the case of Watson v. Fort Worth Bank and Trust). That is, a plaintiff must:

- identify the specific practice being challenged; and
- offer statistical evidence to show that the questioned practice significantly disadvantaged a protected group.

For example, in a layoff case, a plaintiff may show that an employer's college degree requirement for retention in a particular job served to significantly and disproportionately exclude black employees who were otherwise qualified to perform the job. If such statistical

evidence is offered by a plaintiff, the employer must then demonstrate that the degree requirement is a valid predictor of job performance.<sup>1</sup>

On the other hand, disparate treatment claims, involve allegations of intentional discrimination against employees because of their age, race or gender. Accordingly, intent is the deciding issue. In addition to anecdotal testimony, a plaintiff may proffer evidence showing statistically significant differences in employment outcomes between employees argued to be similarly situated. Such statistical evidence alone may lead the decision-maker to an inference of intentional discrimination.

The critical aspect of the statistical evidence in treatment claims is the notion of "similarly-situated" employees. Similarly-situated employees are those who share a number of relevant characteristics (e.g., work in a given department or have a given level of education). It is important to note that the plaintiff need not account for all relevant factors to make a statistical case. Unless the statistical comparisons are obviously inappropriate or misleading, the plaintiff's statistical presentation will serve to shift the burden to the employer. Moreover, it is insufficient for the employer to simply articulate what the omitted factors might be and to hypothesize as to what would happen to the disparity if such factors were considered. Thus, if employee age (for example) is significantly correlated to layoff decisions, the employer must statistically demonstrate that omitted measurable factors (extent and type of experience, educational attainment, performance, etc.) would, in fact, eliminate the statistical discrepancy.

In sum, evidence in disparate impact claims generally is always statistical in nature while statistical evidence for a treatment claim can establish a prima facie case of intentional discrimination. The statistical presentation for an impact claim is typically fairly simple and

---

<sup>1</sup> Intent may arise under an impact theory when, in response to the employer's validity defense, a plaintiff can show that a less adverse, equally valid alternative was available which the employer specifically rejected. The Court may then infer that the choice of the more adverse alternative was made in order to intentionally exclude more minority group members.

straightforward whereas that for a treatment claim tends to be more complicated due to the issue of similarly-situated employees.

### III. A SHORT COURSE ON STATISTICAL EVIDENCE

To understand what constitutes valid statistical evidence, one must first recognize that statistical evidence is indirect evidence from which one may infer the possibility of discrimination. To illustrate the concept of indirect evidence and inferential statistics, consider the following.

Assume that you have engaged in a coin-flipping wager with a stranger where you win one dollar for each of ten coin flips that turns up tails. The stranger flips the coin ten times and reports getting ten heads. Should you accuse the stranger of cheating? Suppose you examine the coin and find that it is two-headed or that you saw a videotape of the flips and they were not all heads. In these situations, you have direct evidence that the stranger cheated (i.e., proof of cheating). Suppose, however, that you have no such direct evidence. The only information you have upon which to base an accusation is the result of ten heads in ten flips. The question, then, is whether this reported pattern of results (statistical evidence) is sufficient to infer that you have been cheated. Note that we use the word "infer" rather than "prove" since one cannot conclude with certainty that cheating has occurred only on the basis of indirect evidence.

The question, of course, is how to decide whether or not the indirect statistical evidence is sufficient to infer that you have been cheated. To determine this, you must first compare the pattern of reported results with what you would have expected from a fair coin-flipping game. Assuming that the coin and the flipping process are fair, you should expect to see five heads in ten flips. However, you will not always observe five heads in ten flips because of the existence of random chance factors. For instance, if the stranger reported getting six heads in ten flips, would you even consider accusing him of cheating, although six heads is one more than expected? Probably not. The odds of observing at least six heads in ten flips of a fair coin are

about one-in-three. That is, this is not a particularly "rare" result. Thus, six heads in ten flips does not constitute statistical evidence of cheating. The disparity (one more head than expected) is reasonably attributed to chance.

This then raises the question as to how divergent from the expected result does a particular result need to be in order to allow us to rule out random chance as the sole cause? The odds of observing ten heads in ten flips of a truly fair coin, based on probability theory, are less than one-in-1,000. It would seem reasonable to rule out chance as the explanation for this result and to accuse the stranger of cheating. Even so, it is possible that the stranger was very, very lucky. Suppose the stranger had reported getting nine heads in ten flips. The odds of this result occurring by chance are about one-in-100. Is this a sufficiently unlikely result to allow us to comfortably accuse the stranger of cheating?

What is needed here is a standard which allows the chance probabilities to be evaluated and which allows us to decide whether or not to accuse the stranger of cheating based on the indirect evidence. In statistical jargon, when the difference between actual and expected results is so divergent as to cause us to believe it is unlikely to have been the result of random chance factors alone, the result is called "statistically significant." In employment discrimination matters, the decision-maker (i.e., the jury or the Court) is free to adopt whatever standard of statistical significance is deemed appropriate. In the discussion so far, the difference between the actual and expected number of heads was converted to a probability value to assess the likelihood of chance occurrence. As an alternative to probability values, statisticians often express the divergence between actual and expected outcomes in units called standard deviations. The larger the difference in standard deviation units, the smaller the probability that the difference is due to random chance factors alone. Courts have traditionally adopted a standard of either two or three standard deviation units<sup>2</sup>. Thus, if the disparity between actual

---

<sup>2</sup> While such statistical significance is necessary for a prima facie disparate impact case, it may not be sufficient. A defendant can argue that the disparity is of no practical significance (a value judgement that the disparity is not large enough to be important).

and expected outcomes exceeds either two or three standard deviations, random chance is ruled out and a valid statistical inference of discrimination may be drawn. Conversely, a smaller standard deviation difference does not allow such a statistical inference.<sup>3</sup> This guideline stems from the jury discrimination case of Castaneda v. Partida [430 U.S. 482 (1977)] in which the Supreme Court stated " . . . if the difference between the expected value and the observed number is greater than two or three standard deviations, then the hypothesis that the jury drawing was random would be suspect . . .". This statistical "rule" was then adopted in the employment discrimination case of Hazelwood School District v. United States [433 U.S. 299. 311 n.17 (1977)] in which the Supreme Court's opinion stated: "a fluctuation of more than two or three units of standard deviation would undercut the hypothesis that decisions were being made randomly with respect to race."

How does all of this help us to analyze a claim that layoff decisions, for example, were not made randomly with respect to employee age? Just as in the coin toss example, we need to do two things:

- (1) compare what actually occurred with what we would expect to have occurred if age played no role in the decision-making process; and
- (2) determine if the difference between actual and expected results is "statistically significant" . . . that is, whether the observed difference may be reasonably attributed to chance factors alone.

The critical task is to determine what is expected from a non-discriminatory process. Often, one simply assumes that every employee has the same likelihood of being selected. Hence, if 60 percent of all employees are "older", we would expect 60 percent of all layoffs to be "older."

---

The EEOC's 80 percent rule is an effort to define practical significance, although the rule has many technical difficulties and acceptance by the courts has been mixed. Using a hiring example, if the minority group hire rate is at least 80 percent of the majority group hire rate, a practically significant disparity is not evident, even though there may be a statistically significant difference in the hire rates.

<sup>3</sup> A judge is likely to apply a more stringent standard for statistical significance (e.g., three standard deviations) in a disparate treatment case than in an impact case, due to the secondary nature of statistical evidence and the intent requirement.

If such an analysis produces a statistically significant difference, does this prove that discrimination has occurred? Absolutely not. We have examined the available indirect evidence and made an inference (based upon our knowledge of the probabilities involved) that it is very unlikely that the process was age-neutral. However, it is possible (although not probable) that chance alone did cause the difference between actual and expected results. More realistically, there are situations where an employee's age appears to be related to the outcome of the decision process even when the process is (in fact) neutral with respect to age.

**IV. TYPICAL STATISTICAL ANALYSIS IN A LITIGATION SETTING: A LAYOFF EXAMPLE**

Let us begin by assuming a firm has 800 employees and business conditions mandate the layoff of 50 employees. The decisions (or recommendations) are made on a business unit level (e.g., by department heads).

A. Plaintiff's Analysis

The plaintiff will usually analyze the data by combining all employees into a single, monolithic group and by comparing the layoff rates of "older" versus "younger" employees. "Older" may be defined in numerous ways. Usually, "older" is defined as being at least forty (40) years old<sup>4</sup>. Another common definition of "older" is "at least as old as the plaintiff". The plaintiff may choose any age break that "works" and which can be justified. Thus, even though no significant disparity may exist between those under and over age forty, you must be concerned with significant disparities which may exist using some other definition of an "older" employee.

Assume that the plaintiff presents the following data:

|           | Number of Employees |                 |                           | Layoff Rate<br>(Percent)<br>$(1)/(3) \times 100$<br>(4) |
|-----------|---------------------|-----------------|---------------------------|---|
|           | Laid Off<br>(1)     | Retained<br>(2) | Total<br>$(1)+(2)$<br>(3) |   |
| "Older"   | 40                  | 360             | 400                       | 10.00%  |
| "Younger" | 10                  | 390             | 400                       | 2.50  |
| TOTAL     | 50                  | 750             | 800                       | 6.25  |

(expected older = 50 layoffs x 50% older = 25)  
(4.24 standard deviations from expected 25)

---

<sup>4</sup> This definition is required for an impact case but alternative definitions may be used for a treatment claim.



If older and younger employees on average share the same risk of a layoff, we would expect an age-neutral process to yield 25 older layoffs (half of the workforce is older . . . thus, we expect that half of all layoffs will be older). As there are more older layoffs (40) than we expect (25), the result is adverse to older employees and we need to determine the likelihood that a truly fair system could produce such an adverse result simply by chance alone. Using a statistical test known as the Yate's Corrected Chi-Square test, we find that this is indeed a fairly rare occurrence (4.24 standard deviations . . . expected to occur by random chance alone only once in more than 40,000 trials).

The above plaintiff's analysis is a "single hat" approach analogous to placing 800 equally-sized colored tickets in a hat (red tickets for "older" employees, green tickets for "younger" employees).

According to the plaintiff's theory, the employer should mix the tickets in the hat and then randomly withdraw fifty (50) tickets. All tickets in the hat are assumed to be competing against one another and, since each ticket is the same size, each ticket is assumed to have the same initial chance of being withdrawn. Under this "model" of the process, an age-neutral selection process should produce a mix of red and green tickets among the 50 selected tickets which is roughly the same as the original mix of red and green tickets. If the proportion of red tickets among the 50 selections is substantially greater than the original proportion of red tickets, plaintiff will claim statistical evidence of age discrimination. To play this employment discrimination "lottery game", the plaintiff need only obtain information concerning the number of red and green tickets in the hat and the number of red and green tickets withdrawn from the hat.

Alternatively, a plaintiff may restrict the analysis to those employees deemed similar to the plaintiff in terms of department, exempt status, job function, etc. In this way, the layoff experience of all other employees is ignored. That is, a single hat is utilized but the number of

tickets considered is restricted. Again, the plaintiff will seek to find a hat which "works" and which can be justified.

B. Defendant's Analysis

The critical aspects concerning the plaintiff's analysis are the existence of a single hat and the assumption that all tickets are equally-sized. In other words, it is assumed that all "older" and "younger" employees were competing against one another for a limited set of positions and, on average, shared the same risk of being laid off (i.e., all employees are similarly-situated). In response, an employer may argue that: (1) the wrong hat is being studied or some tickets were excluded; (2) there are multiple hats; and/or (3) some tickets are larger or smaller than others.

Let us suppose that we discover that there are two types of employees combined in the above who are not competing against each other (i.e., two hats with equally-sized tickets within each):

Managers

|           | <u>Number of Employees</u> |                        |                                | <u>Layoff Rate<br/>(Percent)</u><br>(1)/(3) x 100<br>(4) |
|-----------|----------------------------|------------------------|--------------------------------|--|
|           | <u>Laid Off</u><br>(1)     | <u>Retained</u><br>(2) | <u>Total</u><br>(1)+(2)<br>(3) |  |
| "Older"   | 21                         | 279                    | 300                            | 7.00%  |
| "Younger" | 14                         | 186                    | 200                            | 7.00   |
| TOTAL     | 35                         | 465                    | 500                            | 7.00   |

(expected older = 35 layoffs x 60% older = 21)  
(numerical parity: 0.00 standard deviations)

Support Staff

|           | Number of Employees |                 |                | Layoff Rate          |
|-----------|---------------------|-----------------|----------------|----------------------|
|           | <u>Laid Off</u>     | <u>Retained</u> | <u>Total</u>   | <u>(Percent)</u>     |
|           | (1)                 | (2)             | (1)+(2)<br>(3) | (1)/(3) x 100<br>(4) |
| "Older"   | 5                   | 95              | 100            | 5.00%                |
| "Younger" | 10                  | 190             | 200            | 5.00                 |
| TOTAL     | 15                  | 285             | 300            | 5.00                 |

(expected older = 15 layoffs x 33% older = 5)  
(numerical parity: 0.00 standard deviations)

Clearly, the "single hat" disparity results from:

- the fact that managers were more at risk than support staff (7 percent versus 5 percent);
- the fact that managers tend to be older than support staff (60 percent older versus 33 percent older); and
- the assumption that managers and support staff compete against each other.

To validly rebut the single hat presentation, the employer must be able to defend the notion that managers are not fungible with (do not compete with) support staff.

While this proposition is fairly easy to defend in this example, what if the two hats related to Senior Programmers and Junior Programmers? The functional distinction between these two positions may not be at all clear and the plaintiff will offer a pretext argument, suggesting that Senior Programmers were more at risk because this job group tended to be older. While the same pretext argument can be made concerning managers and support staff, a business necessity response should prevail since the duties of the two job groups are so distinct.

One can also alter the sizes of the tickets in the hats by virtue of key employee characteristics such as education, performance, etc. Such characteristics must also be able to withstand a pretext attack. For example, subjective performance appraisals [or the results of a forced ranking system<sup>5</sup>] might lead to different ticket sizes for otherwise fungible employees and might explain the disparity, but it is very difficult to defend a situation where older employees are viewed (rated) as poorer performers than younger employees. This is consistent with a plaintiff's argument that older employees were subject to age-stereotyping and that biased performance evaluations were the vehicle for age discrimination.

In sum, the employer must statistically demonstrate that studying the relevant hat, accounting for multiple hats and/or accounting for different ticket sizes would make the plaintiff's statistical disparity disappear. However, it cannot be over-emphasized that such a defense requires the employer:

- (1) to justify the appropriate hat or hats for study;
- (2) to justify why tickets are of different sizes; and
- (3) to have the data in order to determine the number of hats and the sizes of the tickets.

---

<sup>5</sup> A forced ranking system rank orders a group of otherwise fungible at-risk employees, assigning lower ranks to the less "at-risk" employees. If three (3) selections must be made from a group of 15 force-ranked employees, the employees ranked "13", "14" and "15" are those to be laid-off. A plaintiff's statistician can examine the age implications of those situations where the forced rankings were not followed and possibly demonstrate an age bias. While exceptions to the forced rankings are inevitable, there should be no pattern related to age. Thus, all situations where a better-ranked employee is terminated rather than a worse-ranked employee (regardless of age) should be double-checked.

## V. WHY DO A STATISTICAL ANALYSIS IN ADVANCE?

The purpose of such a study is to identify potential class action vulnerabilities prior to a lawsuit being filed; to ensure that the data, documentation, and statistical defenses exist to prove any disparities are caused by non-discriminatory factors; and/or to take corrective actions to eliminate disparities.

### Advantages:

- Reduces risks of class action suits and increases likelihood of successfully defending against such suits, if necessary, by:
  - (i) having data, documentation and evidence on hand to defend against any disparity, hence, raising risks to plaintiffs of proceeding as a class,
  - (ii) demonstrating to the jury that you strove to eliminate any potential discrimination in your workplace before being sued,
  - (iii) reducing the chance that plaintiffs' class allegation will be that the challenged common process is excess subjectivity of decentralized decision makers without overall management control or review, and
  - (iv) being able to prove empirically and to document at trial that disparities are caused by non-discriminatory factors, based on data and documentation collected prior to litigation.

Maintenance of contemporaneous documentation of the decision process will facilitate any future defense of the process required. For example, we recommend making clear distinctions between those who terminate during a reduction-in-force process for voluntary versus involuntary reasons<sup>6</sup>. Many involuntary layoffs are preceded by a voluntary program which usually has an early retirement inducement (i.e., tends to attract older employees). Generally, employees who voluntarily opt to terminate via such a program should not be

---

<sup>6</sup> "Voluntary" may include those employees availing themselves of a voluntary early retirement program or those who ask to be terminated to obtain available economic benefits.

included in any statistical analysis of the involuntary program. Moreover, it is not statistically appropriate to include such voluntary program participants in the analysis as not laid-off since the employer presumably did not consider such individuals for the involuntary program (i.e., such employees were not "at-risk"). Counsel will certainly have much to say about the need to obtain releases from voluntary program participants.<sup>7</sup>

It is further recommended that, when working up the proposed list(s) of reduction-in-force actions, documents should not indicate (directly or otherwise) the employee's age, race or gender. When viewed retroactively in litigation, such planning documents may be damaging. This information may easily be brought to bear separately for analysis purposes by persons other than the decision makers.

Finally, such documentation is presumably subject to discovery and eventual production to a plaintiff. If a plaintiff's statistical analysis ignores key aspects of these materials for no apparent good reason, the plaintiff's expert will have a difficult time defending his studies. Moreover, such documents are an invaluable guide to the defense statistician in creating hats and ticket-sizing factors for the rebuttal analysis.

No amount of proactive planning can eliminate the likelihood of a discrimination claim being filed.

However, a second by-product of such a proactive audit is that defenses will be readily available to be supplied to the EEOC (or a state Human Rights agency) when the initial charge is filed. Hopefully, you may be able to obtain a "no cause" finding and discourage the plaintiff's attorney (particularly if working on contingency) from further pursuit of the case or expanding the claim to an allegation of discrimination against an entire class.

An audit can be done of most employment decision processes, such as:

---

<sup>7</sup> More troublesome is the problem of older employees who wish to retire but, due to the economic benefits of a layoff, ask to be terminated. There is nothing wrong with honoring such requests but failure to document that the layoff event was really voluntary will skew the data such that it may appear older employees were being disproportionately selected for layoff.

Hiring (selection and recruitment);  
Initial placement;  
Starting salaries;  
Current compensation;  
Salary growth (merit/promotional);  
Promotions; and  
Involuntary terminations (for cause/RIFs).

## VI. WHAT IS A PROACTIVE STATISTICAL STUDY?

The first task for the analyst performing the proactive study is to understand how the selection process works:

- (1) what factors determine hat membership; and
- (2) what factors determine ticket size.

Hat membership might be defined on the basis of department, job function, location, timing, exempt status, salary grade, decision-maker, etc. Ticket size may be influenced by educational attainment, time-in-grade, productivity measures, etc.

The second task is to make sure that data and documentation verifying the process exists. For all relevant employees, a data base must be compiled reflecting measures for the factors defining hat membership and ticket size.

The third task is to determine what a plaintiff's statistician could produce with the data.

The fourth task is to conduct a defense analysis to see if the disparity (if any) can be validly explained.

The fifth task is to identify potential problem areas at as micro level as possible (i.e., person or decision-maker). For example, individuals whose salaries appear to be out of line must be identified.

The sixth task is to consult with decision-makers to find out why they think disparities are occurring. If data exists, we must test their explanations. If data does not exist, the employer must set up data collection (past and/or future) to test their explanations.

Finally, if disparities still remain that cannot be explained statistically as being the result of non-discriminatory factors, the company must consider modifying its employment process. Thus, a possible seventh task is to adopt on-line monitoring of the key employment process.

## **VII. WHY NOT DO A STATISTICAL ANALYSIS IN ADVANCE?**

While outside counsel can give you advice on how best to protect the confidentiality and discoverability of a statistical audit, one should always consider the possibility that any results will become public or will be discovered by potential plaintiffs.



The worst thing a company can do is to run a statistical audit which identifies potential vulnerabilities and then fail to proceed to either explain them away or correct them. Such an effort creates a prima facie case for the plaintiff and further creates for the plaintiffs an argument to present to the jury that management was aware of its discrimination problems, but rather than correct them, chose to ignore them and dismiss the statistical findings as invalid. Management must then explain to the jury why they did not think such studies revealed discriminatory practices that should have been changed. Then, the question becomes "why did management do the studies in the first place?" . . . a very difficult question to answer.

Even if a company is committed to following through with changes in its practices, if necessary, there still may be a risk of some back exposure. Plaintiffs claiming discrimination may be able to recover damages for company actions in some legally defined period prior to the change.

#### **VIII. TO DO OR NOT TO DO, THAT IS THE QUESTION**

It is my belief that if a company is fully committed to not only doing a full audit, but also fully responding to its findings, the ultimate benefits of armor against future litigation, better documentation of and better understood employment practices, and fairer and more controlled decision making should outweigh the risk of some potential damages for prior actions. Conversely, unless such a total commitment exists, I believe that undertaking such an audit is folly and should be paid for by the plaintiff bar and not by the company, since its benefits will accrue to the plaintiff and not to the company.