IDENTIFYING SIMILARLY SITUATED

EMPLOYEES IN EMPLOYMENT

DISCRIMINATION CASES

SUMMER 1991

Kevin Gilmartin, Ph.D.[†]

I. STATEMENT OF THE PROBLEM

Many, if not most, Title VII[1] employment discrimination cases involve claims of disparate treatment.[2] Statistical proof of disparate treatment requires a comparison of similarly situated members of contrasting groups:

> The key [to statistical evidence of discrimination] is usually comparative evidence. The plaintiff will normally offer a comparison of treatment between protected group members and majority group members, contending that comparably situated protected group members were treated less favorably than majority group members. The defendant will normally contend that the comparisons made by the plaintiff are inappropriate because the individuals are not comparably situated, and/or offer examples of allegedly comparably situated protected group members and majority group members who received similar treatment. In all instances of comparative evidence, the critical question is whether the persons sought to be compared by either the plaintiff or the defendant are in fact in a comparable factual setting.[3]

The intent of this article is to describe some methods for identifying pools of protected group members and nonprotected group members who are "in a comparable factual setting."[4]

Most specific allegations of employment discrimination concern pools of applicants or employees from which selections were made.[5] The allegation is that members of a protected group were selected at lower or higher rates than they should have been (based on the rates at which other persons were selected) and that they suffered some harm from the differential selection rates. Obvious examples of such situations are hirings from among pools of applicants and promotions from among employees at a particular job level. Less obvious examples of selections are the choices of which department new hires should be placed in and which employees should receive training. Being "selected" may also be a negative event for the employee; thus, rates of receiving disciplinary actions and decisions about which employees to lay off are analytically the same kind of situation (except that class members claim that they were overselected rather than underselected).

To perform simple analyses of selection rates to support or refute the allegations, the class members and non-class members being compared must be similarly situated, that is, the pools from which the selections were made should be constructed such that relevant eligibility criteria, qualifications, and circumstances are similar for

the two groups. The more carefully the pools are constructed, the more valid the analyses and conclusions will be.

The alternative to performing simple analyses of carefully constructed pools of similarly situated persons is to perform complex analyses that take differing qualifications and circumstances into account statistically. The most commonly used complex analyses in employment discrimination cases are linear regression analyses (with a continuous outcome, such as salary) and conditional or unconditional logistic regression analyses (with a dichotomous outcome, such as selection versus nonselection). The methods of identifying similarly situated persons described in this article allow the analyst to perform simple and easily understood analyses rather than complex multivariate analyses.[6]

If pools are not predefined by the applicants' or employees' actions (e.g., a pool of employees who have committed a particular infraction), then the analyst will have to construct pools by identifying the persons with similar relevant characteristics. The definition of a pool may depend on several factors. For example, a pool of employees eligible to receive a particular promotion may consist of those persons who: (a) were already hired by the employer as of a certain date; (b) were at a particular job level; (c) were in a particular occupation or had a particular type of experience; (d) were in the organizational unit in which the promotion took place; and (e) had worked sufficient time since being hired or since their previous promotion. As another example, a pool of employees eligible to be sent to a training course may consist of those persons who: (a) were already hired by the employer as of a certain date; (b) were within a range of job levels; (c) were in a particular occupation; (d) were in a particular organizational unit; and (e) had not already had that kind of training. The particular characteristics that determine eligibility for being selected will vary from one kind of selection decision to another and from one employer to the next, and the analyst must become familiar enough with the employer's personnel system to identify those characteristics. However, even after developing a list of eligibility criteria (such as in the examples above), it may be no easy matter to construct the pool of class members and nonclass members from which the selection was made.

To make the task of forming pools of similarly situated persons a bit more difficult, the criteria and even the kinds of variables used to define the pools might change from instance to instance of the same kind of selection decision. For example, some vacancies may be restricted to employees in particular units within the organization and others may be open to qualified employees throughout the organization. Some training may be

restricted to particular job levels, and other training may not. Some training may be only for employees who have not yet had the training, and other training may be open for repeat attendance. So to perform valid analyses of the selection rates of men and women, or blacks and whites, across many separate selection events, it may be necessary to construct the pools in a number of different ways to take into account the differing circumstances of the selection events. This task may seem daunting, but to the extent that the pools are not defined properly, the analyses and any conclusions drawn from them may be invalid.

The following sections present partial solutions to these difficulties based on available data on the composition of the work force (Section II), more valid solutions based on tailoring each pool to match the date and circumstances of its particular selection event (Section III), reasons for possibly needing two or more pools for a single selection event (Section III.A), and computerized methods for constructing the pools from an employee history file (Section III.B).

II. PARTIAL SOLUTIONS

A.   Pools Constructed from Current Work

Force Descriptive Statistics

The easiest method of constructing pools of similarly situated employees from which some kind of selection was made is to use current work force descriptive statistics. Most organizations have available or can easily compile counts of their employees by race, sex, job level, job title, and department within the organization. However, available work force statistics are less likely to be disaggregated by some of the other characteristics that may be relevant to identifying employees eligible to be selected, such as amount of tenure and other experience, date hired, date of last promotion, and previous in-house training. Nevertheless, analyses will be somewhat more valid if pools are defined on the basis of some of the eligibility criteria rather than assuming that all of the employees were eligible for each selection.

Three additional potential problems exist, however, with using current work force counts to estimate the composition of pools of similarly situated employees. All three problems are the result of using a proxy pool of presently available and eligible employees (i.e., the current work force counts) rather than knowing who was in the real pool at the time the selection was actually made. First, if the work force composition has been changing rapidly over time, current work force counts may not be an accurate estimate of the numbers of men and women or black and white employees who were available for selection at some time in the past. (An inaccurate estimate of the composition of a pool can produce an apparent significant disparity in the selection rates between the two groups when an accurate estimate would not have shown a significant disparity, or vice versa.) On the other hand, if the proportion of employees belonging to the protected group (e.g., women, blacks, Hispanics) has not changed much over the time period relevant to the allegations of discrimination, and if their distribution throughout the organization by department, job level, and occupation has not changed much, then current work force counts may be an acceptable estimate of the pools of eligible employees in past years.

The second problem is that standard statistical tests[7] for bias in the selection process are based on the assumption that selections are made from a particular pool, that is, that the selectees were in that pool prior to selection. But if current work force data are used to construct the pools, the persons selected in the past may not be in those pools because they are no longer employees. Or even if they are still employed by the organization, they may now be at a different job level or in a different department and therefore may not be included in the

pool (based on current data) designed to match their selection in the past. The result is that some of the selectees are not included in the pools of employees eligible to be selected.[8]

The third potential problem is that the number of protected or nonprotected group members selected may be greater than the number of such persons in the estimated pool of eligible employees. For example, based on current work force statistics, five women are eligible for a particular promotion, yet over the past year six such women have been promoted; or what seems even more puzzling, the pool of employees eligible to receive a particular kind of training contains no Hispanics, yet last year two Hispanics received such training. The source of this problem is similar to the second problem—the pools are constructed from current work force counts rather than being composed of the employees who were actually on board and eligible at the time a selection was made. But here, the consequences are more serious because a statistical test cannot be performed at all.[9]

If the available amount of time, resources, or data is insufficient to identify similarly situated employees in a more precise way, the use of current work force counts to estimate the composition of the pool of eligible employees will still be more valid and should be given greater weight by the court than analyses that assume that all employees were eligible for every selection. The next two sections, however, will describe increasingly better methods for identifying similarly situated employees.

B. Pools Constructed from Annual Snapshots of the Work Force

A major improvement over using current work force counts to estimate the composition of pools of similarly situated employees is to use a series of work force counts, one for each year during the relevant time period. Many organizations routinely compile such counts at the beginning of each fiscal year or calendar year. As with current work force statistics, the analyst is restricted to use of the particular disaggregations that are available.[10] However, if the composition of the work force has changed appreciably over time, pools based on the composition in the year of the selection will be more accurate estimates of who was available to be selected than pools based on the current composition. Also, if the pools are constructed from data for the particular year when the selection occurred, it is more likely that the selectees will be in the pools of eligible employees, and it is less likely that the number of protected or unprotected persons selected will exceed the number of such persons in the pool.

A refinement to this analysis method is possible if the work force data allow identification of the

individual employees in each pool, not just aggregate counts. One way of avoiding the second and third problems listed above (i.e., selectees not in the pool and more protected or nonprotected persons selected than the number in the pool) is to exclude from the analyses any selectee who is not in the constructed pool of employees eligible for that selection. Thus, the analyzed selections are forced to be true subsets of the pools of similarly situated employees. The analysis is then addressing a slightly different question. Assuming, for example, that the work force data used to construct the pools are from the beginning of each year and the selections being analyzed are promotions, the analytic question becomes: of those employees at a particular grade level (and possibly in a particular department or occupation) at the beginning of the year, which ones were promoted, and is this result at variance with a sex- or race-neutral process? Not included in this analysis are the selections of employees who were ineligible to be selected at the beginning of the year but became eligible and were selected during the course of the year. The analyst should report the proportion of the selections excluded by this refinement of the analysis method. If the proportion is large, the court may decrease the weight that gives to the analysis results.[11]

If quarterly or monthly data on the composition of the work force are available, they can be used to construct pools of similarly situated employees that are more closely tied to the dates of the selections.[12] The more frequent the snapshots, the less likely that selectees will not be in the pools of eligible employees. And if selections that are not in the pool of eligible employees are excluded from the analysis, more frequent snapshots will result in a smaller proportion of the selections being excluded, which in turn would tend to increase the chances of the court's accepting the analysis as reflecting accurately the organization's treatment of the class members.

III. TAILORED POOLS

Given the increasing validity of the analyses when more frequent snapshots of the composition of the work force are used to construct pools of similarly situated employees, the ideal would appear to be daily snapshots, if such were possible. Assuming that the criteria for eligibility are correct and that only eligible employees are ever selected, then pools of eligible employees on the day of the selection should always contain the selected persons, and no selections would need to be excluded from the analysis.

It is only a small step from considering the creation of pools of similarly situated employees that are specifically tailored to each selection event, which in fact is what is now being done in many employment

discrimination cases. Earlier in the history of Title VII litigation when pools were first created to match the particular circumstances surrounding each selection event, each pool had to be constructed by hand[13] for dozens of selection events. The increasing complexity of cases has required litigants to analyze hundreds or even thousands of selection events, and it has become infeasible to tailor each pool to each selection event by hand. However, when the process of defining how each pool should be constructed is itself computerized, tailoring the pools to the selection events is quite feasible.

A. Multiple Pools per Selection Event

In some situations, it may be necessary to construct two or more pools per selection event. Simple analyses that do not take qualifications and other characteristics into account are based on the assumption that class members and nonclass members in a pool are similarly situated. However, if there is a wide range of qualifications among eligible persons, or if there are discretely different levels of qualifications, and if the protected and nonprotected groups are distributed across the qualification levels differently, then analyses based on a single pool for the selection event may be inaccurate and misleading.

For example, sometimes employees at a particular job level are considered for vacancies at that same job level (e.g., a lateral transfer within the organization or a move to a different kind of job but one at the same job level). They would generally be at a competitive advantage over employees currently at a lower job level (who would tend to have less experience). In this situation, analysis results would be more valid if two pools of similarly situated employees were created: a pool of eligible persons a level below the vacancy and a pool of eligible persons at the level of the vacancy. Similarly, although employees throughout an organization may be eligible for a vacancy, employees already working in the department where the vacancy occurs may have a competitive advantage, in which case separate pools should be constructed of eligible persons in the department and eligible persons outside the department.

Generally, the constructed pools should consist of persons who are approximately similarly situated. It is not necessary that everyone in a pool have identical qualifications, but the results could be misleading if the class members and nonclass members in the pool have differing distributions of qualifications, such that one group tends to be more qualified than the other. When this occurs, the analyst should include either additional eligibility criteria in the definition of the pool so that the remaining protected and nonprotected group members

are more equally qualified on the average, or else create separate pools for two or more qualification levels such that protected and nonprotected group members are approximately equally qualified within each pool.

B.   Computerized Methods for Constructing the Pools

The analyst needs to create or have access to an *employee history file*[14] covering the relevant work force[15] and time period[16] and containing information on all of the variables to be used to construct the pools. An employee history file is a transactional file (i.e., with varying numbers of records per employee) with a record for each major[17] personnel action that affected the status of each employee: being hired, being promoted, receiving a salary increase, being reassigned to another department, and departing from the work force, for example. Each record specifies the employee's status following the personnel action, both on the variables that were changed by the personnel action and on the variables that remained unchanged. Computer programs can be written to extract from the employee history file the status of any employee on any day. In particular, an employee history file can be used to identify all of the employees who met particular eligibility criteria on the day that a selection was made, and such pools of similarly situated employees can be constructed for every selection event that occurred.

If there are relatively few selection events to be analyzed, one can write computer code for constructing each pool of eligible employees. The portion of the computer program that defines one of the pools might be as follows (written here in quasi-English for the sake of the example), and similar computer code could be written for each selection event:

IF THIS IS THE FIRST TRANSACTION FOR AN EMPLOYEE

AFTER THE PARTICULAR SELECTION DATE,

AND IF THE EMPLOYEE HAS NOT JUST BEEN HIRED,

AND IF ACCORDING TO THE PREVIOUS TRANSACTION FOR THIS

EMPLOYEE SHE WAS AT THE APPROPRIATE JOB LEVEL

AND WAS IN THE APPROPRIATE OCCUPATION

AND WAS IN THE APPROPRIATE DEPARTMENT,

THEN OUTPUT A RECORD INDICATING THAT SHE WAS

AVAILABLE AND ELIGIBLE TO BE SELECTED.[18]

If there were hundreds or thousands of selection events (which could be the situation if all of the promotions in a large company over a decade are being analyzed[19]), then writing computer code by hand to identify the available and eligible employees for each selection would not be feasible (or at least would not be efficient). Instead, the analyst can write a computer program that will automatically produce the computer code needed to identify the eligible employees for each selection event. This requires a computerized data file with information about each of the selection events: the date, job level, occupation, department, or whatever other eligibility criteria applied. A computer program can be written to access the information from this file for one selection event at a time and to output appropriate computer code for identifying all of the eligible employees for that selection (e.g., computer code like the example in footnote 18). These hundreds or thousands of lines of computer code can then be accessed by another program that executes them while passing through an employee history file once, thereby creating all of the needed, specially tailored pools of similarly situated employees in a single step![20]

Once the computer program has constructed the pools, determined the number of class and nonclass members in each pool, and identified the persons who were selected, then analyses are relatively straightforward.[21] If two or more pools exist for a single selection event (because there were two or more sets of employees with distinctly different qualifications), the analyst will have to compute the outcome for class members within each pool and then aggregate the results properly across the two or more pools to determine the amount of disparity for that selection event.[22] Similarly, the analyst wil have to aggregate the results for individual selection events by job level or department (or whatever other way might be meaningful) and over all of the selection events to determine whether there are statistically significant disparities for parts or all of the work force. To the extent pools have been constructed to contain employees who are similarly situated in the eyes of the law, these analyses will provide valid measures of the employer's relative treatment of class members.

IV. CONCLUSION

The intent of most employment equity analyses is to determine what the treatment of a protected group of employees would have been in the absence of discrimination. To be valid, those analyses have to take into account any legally relevant differences between the protected employees and a comparison group of employees (e.g., all of the other employees, white employees, males, or white males). The analyst can take into account differences in qualifications, productivity, and other circumstances either by using methods that statistically control for such variables (such as logistic regression analysis) or by identifying pools of similarly situated employees with minimal average differences between the protected and nonprotected employees within each pool.

In order of increasing accuracy and validity, pools can be based on: (a) the current composition of the work force; (b) the composition of the work force annually during the relevant period; (c) even more frequent snapshots of the composition of the work force; or (d) the specific circumstances of each selection event and the particular employees who met those criteria on that selection day. Pools that are specifically tailored to each selection event are derived from an employee history file by using a computer program that passes through the file and identifies everyone who met the relevant criteria on the date of the selection event. For a few dozen selection events, the computer code defining each pool can be written by hand, but for hundreds or thousands of selection events, one needs a preceding computer program that will produce the computer code that passes through the employee history file.

The advantage of identifying pools of similarly situated persons is that the analyses to detect the effects of employment discrimination can then be simple to perform and simple to explain to the court (although multivariate analysis methods may be more powerful when the qualifications are continuous variables). As with any analysis that is performed many times for subsets of the work force, the results from separate pools need to be aggregated properly to determine whether there is a pattern and practice of employment discrimination that is not detectable in the individual analyses.

NOTES

† Senior Research Fellow, Director of the Litigation Support Program, Director of the Applied Behavioral Sciences Group, American Institutes for Research, P.O. Box 1113, Palo Alto, CA 94302, (650) 493-3550; testifying expert on employment discrimination analyses.

1. 42 U.S.C. §§ 2000e–2000e-17 (1982) (originally enacted as Pub. L. No. 88-352, §§ 701–718, 78 Stat. 253 (1964)).

2. "'Disparate treatment' . . . is the most easily understood type of discrimination. The employer simply treats some people less favorably than others because of their race, color, religion, sex, or national origin. . . . Undoubtedly disparate treatment was the most obvious evil Congress had in mind when it enacted Title VII." Teamsters v. United States, 431 U.S. 324, 335 n.15. *See* B.L. SCHLEI & P. GROSSMAN, EMPLOYMENT DISCRIMINATION LAW 13–22 (1983).

3. *Id.* at 15.

4. In some situations, such pools are *predetermined,* and no special procedures are needed to identify them. For example, if the pools are composed of male and female employees who have applied for particular vacancies within their current organization, an analyst might tentatively accept the applicants in each pool as being comparable—at least in an initial analysis before considering whether refinements are needed—since they have self-selected themselves and have shown sufficient interest in the vacancy to apply and therefore might be presumed to be reasonably similar in terms of circumstances and qualifications. (It is not necessary to assume that all of the protected and nonprotected group members have identical qualifications, only that the distribution of relevant qualifications is the same in each group). The discussion below will focus on more difficult situations where the pools to be compared are not predetermined or self-selected.

Also, in some situations the comparison is to a population *external* to the organization (e.g., a comparison of the proportion of blacks hired with the proportion of blacks in the qualified labor force in the appropriate geographical area.) A number of publications have addressed issues concerning the proper use of Census data to compare the composition of an organization with an external referent. *See, e.g.,* W.B. CONNOLLY, JR., D.W. PETERSON, & M.J. CONNOLLY, USE OF STATISTICS IN EQUAL EMPLOYMENT OPPORTUNITY LITIGATION, § 4.02 (1989); SCHLEI & GROSSMAN, *supra* note 2, at 1351–58, 1361–64 and 1989 supplement at 502–05; D.C. BALDUS & J.W.L. COLE, STATISTICAL PROOF OF DISCRIMINATION § 4.2 (1980). In contrast, the discussion below will focus on identifying pools of similarly situated employees within the organization.

5. Two other types of allegations will not be discussed here: (1) wage or salary discrimination (unequal pay for the same work, similar work, or work of comparable worth to the employer); and (2) discriminatory assignment of performance ratings.

6. Some courts are acquainted with multivariate methods and may even expect them to be presented in an employment discrimination case, although that is currently more often true of linear regression analysis than logistic regression analysis. Multivariate methods have the advantage that they are more powerful statistically if the qualifications are continuous (e.g., amount of tenure) rather than discrete (e.g., being in a particular occupation or department), because some information is lost when a continuous variable is used to form pools of similarly situated employees. On the other hand, the multivariate method may be based on assumptions concerning the statistical properties of the data or the relation of the qualification to the probability of being selected that could be attacked by an opposing expert.

7. For example, Fisher's exact test is appropriate in the situation of selection without replacement (i.e., if multiple selections are made from the pool, the same person cannot be selected more than once from that pool).

8. The argument could be made that the pools are being constructed merely to estimate the proportion or the number of class members who were eligible to be selected, not to identify the particular persons who were eligible, and therefore it does not matter whether the person selected was in the constructed pool or not. For example, if the pool of eligible employees is estimated to have been half male and half female, one would expect

that there was a 50% chance that a woman would be selected, regardless of whether the actual person selected happened to be in the constructed pool. Analytically this is a reasonable argument, but there may be difficulties in explaining this point of view to the court, and at the very least not having the selectees in the pool of eligibles is an apparent weakness in the procedure that may be exploited by the opposing attorney.

9. Part of the statistical test would be equivalent to asking what the probability is of selecting two Hispanics from a pool that contains none, an impossible situation, or of selecting six women from a pool containing only five women where no woman can be selected more than once, which is equally impossible.

10. In fact, fewer types of disaggregations (e.g., by race, sex, job level, job title, department) may be available for past years than could be produced, if requested, for the current work force.

11. Thus the analyst in this situation is faced with the tactical choice of: (a) including all of the selections in the analysis and then having to explain why some selectees are not in the estimated pool of eligibles for that selection; or (b) restricting the analyses to selections that are strict subsets of the estimated pools of eligibles and then having to explain why some selections are not included in the analysis.

12. Personnel departments probably do not have quarterly work force counts already available, and monthly snapshots of the work force are almost certainly not available. However, if the personnel system is computerized and is complete (including information on all of the past employees who have since left the organization), then possibly such counts could be produced upon request. As an alternative, some information may be available from the payroll department about the status of every employee in each pay period during past years.

13. That is, by separately writing the computer code to produce each pool.

14. *See* Bessey, *Employment History File: Essential Element in Employment Discrimination Cases,* (1991) (manuscript submitted for publication).

15. In employment discrimination class actions, the class is often certified to consist of only a portion of the employer's work force (e.g., certain occupations, a range of job levels, or certain departments). It may then not be necessary for the employee history file to contain information on every employee. However, the employee history file will be used to construct pools of employees eligible for each selection event, and if some employees outside the relevant portion of the work force would have been eligible for some selections, they need to be included in the employee history file as well. For example, if the class consists of nonsupervisory positions in the Marketing Department, yet some employees outside the Marketing Department would have been eligible to be selected for vacancies in the department, they, too, should be included in the employee history file.

16. Similarly, the employee history file does not need to contain information about the status of employees prior to the relevant time period as defined by the court unless earlier information is needed to construct the pools. For example, if, according to the employer's personnel procedures, employees must be at a job level for a year before becoming eligible for a promotion to a higher job level, then the employee history file may need to extend back one year prior to the first promotion, in order to be able to distinguish which employees had spent sufficient time at the next lower job level to be eligible.

17. Which personnel actions are "major" depends on the allegations and planned analyses in the case. For example, for the purposes of many lawsuits, monetary awards for exceptional performance might not be important enough to include in the file. However, if the allegations included discriminatory distribution of awards, or if receipt of awards is considered when deciding which employees should be promoted, then awards should be included in the file.

18. In the commonly used SAS® programming language, this program code would look more like the following (with some example values filled in for the selection date, job level, occupation, and department):

```
IF CUR_DATE GT '16JUL89'D AND PREV_DAT LE '16JUL89'D
```

```
AND ACTION NE 'HIRE' AND PREV_LEV = 7
AND PREV_OCC = 'PROGRAMMER'
AND PREV_DEP = 'MARKETING' THEN OUTPUT;
```

(SAS program is a product of SAS Institute, Inc., Cary, N.C.).

19. For example, in one of our recent employment discrimination cases, the relevant work force consisted of 9,000 employees, and the analyses covered a 12-year period.

20. Example programs in the SAS® language for: (a) reading a file of information about selection events and producing computer code to identify the eligible employees; and then (b) executing that computer code to create the pools of similarly situated employees are available by writing or telephoning the author.

21. If the employees in a pool could not be selected more than once (which is referred to as "selection *without* replacement"), then Fisher's exact test based on the hypergeometric probability distribution would usually be used. In some less common situations, employees can be selected two or more times from the same pool (selection *with* replacement)—for example, awards or disciplinary actions received during a year. In this situation, the binomial test based on the binomial probability distribution would be used instead. When a separate pool is created for each selection event, however, it is unlikely that selection with replacement is the appropriate statistical model.

22. The Mantel-Haenszel test is an *approximate* method for aggregating results across independent selection events and has been accepted by courts in a number of employment discrimination cases. *See* Aickin, *Issues and Methods in Discrimination Statistics,* in STATISTICAL METHODS IN DISCRIMINATION LITIGATION 159, 198–200, 202 (D.H. Kaye & M. Aickin eds. 1986); J.L. GASTWIRTH, STATISTICAL REASONING IN LAW AND PUBLIC POLICY: VOLUME 1, STATISTICAL CONCEPTS AND ISSUES OF FAIRNESS, 230–33 (1988).

 An *exact* method for aggregating such results was first developed for employment discrimination lawsuits by Dr. Stephan Michelson under the name "multiple pools exact test." The exact aggregation of results involves: (1) computing the probability of every possible outcome from the selections from one pool (e.g., the probabilities of zero, one, or two women being selected when two random selections are made from the pool); (2) computing the probability of every possible outcome from the selections from the second pool; (3) multiplying the probability of each outcome from the first pool by the probability of each outcome from the second pool; and then (4) adding together all of the products that result in the same number of total selections across the two pools (e.g., two and zero selections of females from the two pools, one and one selection, and zero and two selections all result in a total of two selections). This process of aggregating two probability distributions under an assumption of independence is technically called "convolving," and the process can be applied repeatedly to any number of analysis results.

 This exact method of analyzing selection events and then aggregating the results across pools is embodied in a computer program called MULTEVENT, which was written by staff at the American Institutes for Research under a contract from the Office of the General Counsel, Department of the Navy. Versions of MULTEVENT designed for use on IBM mainframe computers or on MS-DOS microcomputers, along with a user's manual and a validation report (which contains some simple examples of the aggregation process that can be performed on a hand calculator), are available by writing or telephoning the author.

 Employment discrimination cases in which either the approximate or the exact method has been used include: Dees v. Orr, Civ. S-82-471 LKK (E.D. Cal. 1988); Harrison v. Lewis, C.A. 79-1816 (D.D.C. 1982); Hogan v. Pierce, 31 FEP 115 (D.D.C. 1983); Moore v. Naval Air Rework Facility, C-81-0905 MHP (N.D. Cal. 1984); Trout v. Garrett, C.A. 73-55 (D.D.C. 1990); Williams v. Secretary of the Air Force, C.A. C79-397-TV (W.D. Wash. 1989); Charles v. Naval Supply Center, EEOC # 033-081-X1180, 091-85-X0016, X0017, & X0018 (San Francisco Office 1986); Sylvester v. Secretary of the Air Force, EEOC # 091-88-X1821 (San Francisco Office 1990); Lewis v. Brady, C.A. 82-0918 TFH (D.D.C. 1990); Grimes v. General Services Admin., EEOC # 091-86-X1084 (San Francisco Office 1988); Johnson v. Garrett, C.A. 73-702-CIV-J-M (M.D. Fla. 1991).