

June 1988

Measures of Employment Discrimination: A Statistical Alternative to the Four-Fifths Rule

Marion Gross Sobol

Charles J. Ellard

Follow this and additional works at: <https://scholarship.law.berkeley.edu/bjell>

Recommended Citation

Marion Gross Sobol and Charles J. Ellard, *Measures of Employment Discrimination: A Statistical Alternative to the Four-Fifths Rule*, 10 BERKELEY J. EMP. & LAB. L. 381 (1988).

Link to publisher version (DOI)

<https://doi.org/10.15779/Z38B05J>

This Article is brought to you for free and open access by the Law Journals and Related Materials at Berkeley Law Scholarship Repository. It has been accepted for inclusion in Berkeley Journal of Employment & Labor Law by an authorized administrator of Berkeley Law Scholarship Repository. For more information, please contact jcera@law.berkeley.edu.

Measures of Employment Discrimination: A Statistical Alternative to the Four-Fifths Rule

Marion Gross Sobol†
Charles J. Ellard††

The EEOC's Uniform Guidelines suggest a "four-fifths" test as a criterion for adverse impact in age, ethnic, race and sex discrimination cases. The use of the rule is explained and demonstrated and the results are compared to those that would be obtained by using a binomial test which allows probability statements and reflects differences in sample size. The authors develop a "crossover point" where the results obtained from the four-fifths rule and the binomial test are equal and discusses those cases where there is divergence. The analysis shows that the four-fifths rule is less reliable in sex discrimination cases than in race discrimination cases. Moreover, for small numbers of hires the four-fifths rule is more demanding of small employers than a statistical inference criterion would be.

A multitude of ratios and indexes have been used to indicate racial, ethnic, sex and age discrimination.¹ One ratio, the "four-fifths rule," has gained official government sanction.² This rule, which is essentially a

† Chairman, Management Information Sciences, Cox School of Business, Southern Methodist University; B.A., 1951 Syracuse University; M.A., 1958, Ph.D., 1961, University of Michigan; Member, Phi Beta Kappa; Vice President, Decision Sciences Institute; Book Review Editor, *INTERFACES*; Associate Editor, *DECISION SCIENCES*.

†† Professor of Economics, Pan American University; B.S. 1958, University of Connecticut; M.S. 1961, University of Rhode Island; Ph.D., 1974, University of Houston; Member, American, Southern and Western Economics Associations.

1. Comment, *Statistics and Title VII Proof: Prima Facie Case and Rebuttal*, 15 *HOUS. L. REV.* 1030, 1052 (1978) (methods of statistical analysis used in Title VII actions include independent proportions, multiple regression and analysis of variance, a percentage allowance test, the four-fifths test and the binomial distribution).

2. "A selection rate for any race, sex, or ethnic group which is less than four-fifths . . . of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact . . ." Uniform Guidelines on Employee Selection Procedures (1978), 29 C.F.R. § 1607.40 (1987); see also Adoption of Questions and Answers to Clarify and Provide Common Interpretation of the Uniform Guidelines on Employee Selection Procedures, 44 Fed. Reg. 11,996 (1979).

The circuits have differed on the weight which they have ascribed to the EEOC guidelines. The Eighth Circuit has applied them as if they were mandatory, requiring strict compliance. Firefighters

© Industrial Relations Law Journal, Volume 10, No. 3, 1988.

"rule of thumb"³ rather than a statistical criterion, is used because it is simple to calculate and to understand. For the same reasons that it is simple to use, however, in certain circumstances it is an unreliable indicator of discrimination. A statistical significance test for the binomial distribution ("binomial test") is preferable because it provides a statistical probability criterion which estimates the probability of occurrence of certain events and accounts for sample size. Thus, probability estimates can be made which are not possible using an arbitrary rule of thumb. This Article will show under what circumstances the four-fifths rule and the binomial test give the same results and under what circumstances they give different ones.

The paper is divided into three Parts. The first Part considers theories of discrimination with particular attention to the type of ratios used as statistical evidence. Next, the four-fifths rule is demonstrated, and finally, the four-fifths rule is compared to the binomial test.

I

THEORIES OF DISCRIMINATION

While courts apply the Civil Rights Act of 1964 to federal social programs largely through Title VI⁴ and certain additional executive orders,⁵ courts have issued the more important decisions with respect to Title VII.⁶ Thus, it is useful to look at the history of Title VII litigation insofar as it bears on problems of statistical criteria in cases relating to discrimination. This section will examine some of the statistical problems in arranging proof to make a *prima facie* case of discrimination under Title VII.

Inst. for Racial Equality v. City of St. Louis, 616 F.2d 350, 355-56 (8th Cir. 1980). The Fifth Circuit has held that the guidelines merely provide "a valid framework" for assessment, not necessarily requiring mandatory compliance. *United States v. Georgia Power Co.*, 474 F.2d 906, 913 (5th Cir. 1973). The Ninth Circuit has held that the rule does not require strict adherence. *Clady v. County of Los Angeles*, 770 F.2d 1421, 1428 (9th Cir. 1985) ("[t]his circuit analyzes whether the statistical disparity is 'substantial' or 'significant' in a given case," *id.* at 1428-29); *see also* *Guardians Ass'n of New York City Police Dept., Inc. v. Civil Serv. Comm'n*, 630 F.2d 79, 90-91 (2d Cir. 1980). *See generally* B. SCHLEI & P. GROSSMAN, *EMPLOYMENT DISCRIMINATION LAW* 92-97 (2d Ed. 1983), 324-25 (2d Ed. 1985 Supp.); 3 LARSON, *EMPLOYMENT DISCRIMINATION* § 74.52 (1987).

3. Eighty percent is merely a "rule of thumb": "Smaller differences in selection rate may nevertheless constitute adverse impact, where they are significant in both statistical and practical terms or where a user's actions have discouraged applicants disproportionately on grounds of race, sex, or ethnic group." 29 C.F.R. § 1607.4D (1987).

4. 42 U.S.C. §§ 2000d to 2000d-4 (1982).

5. *See* Exec. Order No. 12,250, 45 Fed. Reg. 72,995 (1980).

6. 42 U.S.C. §§ 2000e to 2000e-17 (1982). Title VII provides, in part, that it is an unlawful employment practice for an employer "to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin." 42 U.S.C. § 2000e-2(a)(1) (1982).

Courts have recognized four general categories of discrimination.⁷ The first category, disparate treatment, is the most easily understood, and represents what is commonly seen as the primary target of the Civil Rights Act of 1964 at the time of its enactment.⁸ Courts find disparate treatment where equals are treated unequally or unequals are treated equally.⁹ Examples include refusing to consider blacks for employment, paying a woman a lower wage than that paid a man for the same work, and discharging a Hispanic worker for an offense for which whites are given lesser discipline. A landmark disparate treatment case, *McDonnell Douglas Corp. v. Green*,¹⁰ dealt with the proper order and nature of proof in actions under Title VII. This case involved the discharge of a black long-time employee of the McDonnell Douglas Corporation, who claimed he had been discharged illegally in 1964, and also not rehired in 1965, because of his involvement in civil rights activities and because of his race. The Court held that while the employer's reason for discharging the plaintiff sufficed to meet the prima facie case, the plaintiff must be afforded a fair opportunity to show that this proffered reason "was in fact pretext."¹¹

Under the second category of discrimination, there is a challenge to policies or practices which perpetuate in the present the effects of past discrimination. Present effects of past discrimination are exposed in cases where blatant discrimination had existed before the passage of the Act, and upon passage of the Act, company policies were "jerry-built" to give the appearance of compliance; such policies serve to perpetuate historic discrimination.

An important early case of present effects of past discrimination is *Quarles v. Philip Morris, Inc.*¹² The court was asked to consider whether restrictive departmental transfer and seniority policies "[were] intentional, unlawful employment practices because they [were] superimposed on a departmental structure which was originally organized on a racially segregated basis."¹³ Prior to the effective date of the Civil Rights Act of 1964, the employer employed blacks, but only in the least desirable departments. Upon passage of the Act, the company ceased this practice,

7. See generally B. SCHLEI & P. GROSSMAN, *supra* note 2, at 1-12.

8. In S. Rep. No. 1137, 91 Cong., 2d Sess. 4 (1970), the Senate Committee on Labor and Public Welfare stated:

In 1964, employment discrimination tended to be viewed as a series of isolated and distinguishable events, for the most part due to ill-will on the part of some identifiable individual or organization.

B. SCHLEI & P. GROSSMAN, *supra* note 2, at 1 n.6.

9. B. SCHLEI & P. GROSSMAN, *supra* note 2, at 15.

10. 411 U.S. 792 (1973).

11. *Id.* at 804.

12. 279 F. Supp. 505 (E.D. Va. 1968); see also *NLRB v. Local 269, IBEW*, 357 F.2d 51, 55-56 (3d Cir. 1966). Compare *Whitfield v. United Steelworkers*, 263 F.2d 546 (5th Cir. 1959).

13. 279 F. Supp. at 510.

but subsequently flatly barred transfers between departments, or required that blacks forfeit their seniority if they wished to transfer to a different and higher paying department. These practices tended to lock blacks into the department in which they had been originally placed. The federal district court ruled against Philip Morris, stating that it had "intentionally engaged in unlawful employment practices by discriminating on the basis of race against Quarles."¹⁴

The third category, adverse impact, is found in cases where facially neutral employment policies or practices impact one group more harshly than another "and cannot be justified by business necessity."¹⁵ An example is the use of a general intelligence test (as a hiring prerequisite) which disqualifies substantially more blacks than whites and which cannot be shown to be job related; that is, it does not predict successful job performance. The landmark case of adverse impact was *Griggs v. Duke Power Co.*¹⁶ This case involved the company's practice of requiring job applicants to take a general intelligence test and its hiring requirement that the applicant be a high school graduate. These requirements were applied to new hires and transfers in labor and coal handling positions—jobs which the court ruled did not warrant such requirements.¹⁷ In this landmark case, the Supreme Court ruled that Title VII prohibits "not only overt discrimination but also practices that are fair in form, but discriminatory in operation."¹⁸ "[G]ood intent or absence of discriminatory intent" will not redeem a procedure or mechanism which operates to exclude minority groups.¹⁹ As a consequence, adverse impact cases rely more heavily on statistics for the proof needed to establish a prima facie case of discrimination²⁰ than do other categories because the court is looking for substantially disparate effects rather than discriminatory motivation.

The final category of discrimination, reasonable accommodation, involves instances where employers fail to make reasonable accommoda-

14. *Id.* at 519.

15. *International Bhd. of Teamsters v. United States*, 431 U.S. 324, 335-36 n.15 (1977).

16. 401 U.S. 424 (1971).

17. *Id.* at 431.

18. *Id.*

19. *Id.* at 432.

20. *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975) (the plaintiff establishes a prima facie case by showing "that the tests in question select applicants for hire or promotion in a racial pattern significantly different from that of the pool of applicants").

Testing and screening have been scrutinized rather carefully in adverse impact cases. These cases may be classified according to their use of scored tests, such as I.Q. tests; nonscored objective criteria, such as educational qualifications; arrest or conviction records or garnishment incidents; or subjective criteria such as appraisal of interest, aptitude, personality, ability to fit in, aggressiveness and leadership. Questions of validating these criteria, which are used in the hiring and promotion process, are important but are beyond the scope of this paper.

tion to an employee's handicap²¹ or religious practices and preferences.²² *Cummins v. Parker Seal Co.*²³ is a typical case involving work-schedule accommodations in which the worker claimed discriminatory discharge because his religion prohibited him from working on Saturdays and other holy days. The employer was unable to show that undue hardship would be involved in accommodating its business to the employee's religious practices.

The use of statistics in establishing a prima facie case of discrimination is particularly crucial because the efficacy of such statistics often determines whether the court will hear the case or not.²⁴ Statistics are frequently influential in creating an inference of discrimination; if accepted, the statistics can establish the prima facie case of discrimination and shift the burden of proof to the defense.²⁵ The utility of statistical evidence will depend upon the type and purpose of the evidence in proving the particular category of discrimination: it may be useful but not sufficient in disparate treatment cases and indispensable in adverse impact cases where plaintiffs must establish that facially neutral policies or practices impact on one group more severely than on another.²⁶

In adverse impact cases, two kinds of statistical comparisons can be used: pass/fail ratios and population/work force comparisons.²⁷ In pass/fail rates, the comparisons are between the relative success of minority applicants compared with whites. For example:

21. Rehabilitation Act of 1973, 29 U.S.C. §§ 701-796 (1982).

22. Section 701(j) of Title VII permits discrimination if an employer demonstrates that he is unable to reasonably accommodate to an employee's or prospective employee's religious observance or practice without undue hardship on the conduct of the employer's business. 42 U.S.C. § 2000e(j) (1982). The Supreme Court has held that accommodations which require an employer to bear more than de minimis cost are an undue hardship. *Trans World Airlines v. Hardison*, 432 U.S. 63, 84 (1977).

23. 516 F.2d 544 (6th Cir. 1975), *aff'd*, 429 U.S. 65, *but vacated*, 433 U.S. 903 (1977) (decision overruled in light of *TWA v. Hardison*).

24. More than one judge has expressed anguish at the statistical nature of contemporary employment discrimination law. One judge admitted that he struggled to sift through all the expert explanations and make a credibility determination while "not highly trained in statistical techniques." *Dickerson v. United States Steel Corp.*, 472 F. Supp. 1304, 1314 (E.D. Pa. 1978), *vacated*, *Worthy v. United States Steel Corp.*, 616 F.2d 698 (1980). Another remarked that "Title VII cases are contests between college professor statisticians who revel in discoursing about advanced statistical theory." *Otero v. Mesa County Valley School Dist. No. 51*, 470 F. Supp. 326, 331 (D. Colo. 1979).

25. While blatant disparities may be sufficient to establish a prima facie case, a lesser showing will require plaintiff to show a causal nexus between the challenged practice and resultant disparities in order to create an inference of unlawful discrimination. *Carroll v. Sears, Roebuck & Co.*, 708 F.2d 183 (5th Cir. 1983).

26. Lamber, Reskin & Dworkin, *The Relevance of Statistics to Prove Discrimination: A Typology*, 34 HASTINGS L.J. 553, 576-90 (1983).

27. B. SCHLEI & P. GROSSMAN, *supra* note 2, at 317-19 (other types of statistical proof of adverse impact include regression analysis, employer's overall hiring practices in filling upper-level positions and cohort analyses in promotion and pay cases).

$\frac{\text{Total number of successful black applicants}}{\text{Total number of black applicants}}$	compared to	$\frac{\text{Total number of successful white applicants}}{\text{Total number of white applicants}}$
--	-------------	--

In the population/work force ratios, the comparison is between the percentage of some minority of the relevant population²⁸ and the percentage of the same minority in the employer's work force. For example:

$\frac{\text{Total number of blacks employed by the employer}}{\text{Total number of persons employed by employer}}$	compared to	$\frac{\text{Total number of blacks in the population or work force in the relevant geographical area}}{\text{Total number of persons in the population or work force in the relevant geographical area}}$
--	-------------	--

The question remains as to what weight should be given to statistical evidence in proving discrimination in employment. Judges are asked to determine whether the statistical evidence reveals a substantial disparity between protected groups.²⁹ The Supreme Court has addressed the issue of "substantial disparity" in two recent cases: *Castaneda v. Partida*³⁰ and *Hazelwood School District v. United States*.³¹ In both cases, the Court pointed toward the use of a precise statistical measure, the standard deviation, as the method to appropriately gauge the significance of observed disparities:

As a general rule for large samples, if the difference between the expected value and the observed number is greater than two or three standard deviations, then the hypothesis that [the disparity was random] would be

28. A threshold inquiry into the relevant population upon which to base comparisons is crucial and often determinative. When the court is looking at a neutral device which allegedly has disparate impact, it may allow a large geographical scope in statistics. See, e.g., *Dothard v. Rawlinson*, 433 U.S. 321 (1977) (height requirement: Supreme Court used statistical evidence derived from the entire United States); *Johnson v. Goodyear Tire & Rubber Co.*, 491 F.2d 1364 (5th Cir. 1974) (high school certificate: court used statistical evidence from the State of Texas). When the plaintiff relies on statistical underrepresentation itself, the court will limit the statistical base to those individuals who are potential employees and who live within a reasonable recruiting area. See, e.g., *Hazelwood School Dist. v. United States*, 433 U.S. 299 (1977). See generally, Smith & Abram, *Quantitative Analysis and Proof of Employment Discrimination*, 1981 U. ILL. L.F. 33, 59-62.

29. According to the EEOC, statistical significance of a disparity ordinarily means that the relationship should be "sufficiently high so as to have a probability of no more than one (1) in twenty (20) to have occurred by chance." 29 C.F.R. § 1607.14B(5) (1987).

30. 430 U.S. 482 (1977). While this case involved discrimination in grand jury selection under the Texas "key man" system, its statistical analysis has been reaffirmed in employment discrimination cases under Title VII. See *infra* note 32 and accompanying text.

31. 433 U.S. 299 (1977).

suspect.³²

The Supreme Court has ruled that "where gross statistical disparities can be shown, these alone may in a proper case constitute *prima facie* proof of a pattern or practice of discrimination."³³

In the following sections, we will discuss when it would be preferable to utilize a binomial test instead of the four-fifths rule to indicate discrimination. The discussion is based on a mathematical derivation of what would happen using the two criteria. The binomial test, which allows for probabilities of the occurrence of different events by chance, is preferable to a "rule of thumb" in the analysis of statistical data. While legal scholars have discussed the importance of using statistical criteria,³⁴ no study has been made of the cases where the results of the four-fifths rule and the binomial test give different results. We have derived a "crossover point" where the results are equal and will look at the cases where there is divergence. Thus, it will be possible to cite situations where the four-fifths rule will be satisfactory and cases where the binomial test should be employed.

II

DETERMINATION OF ADVERSE IMPACT USING THE FOUR-FIFTHS RULE

The Uniform Guidelines provide a basis for determining when an employment selection process adversely affects the opportunities of a race, sex or ethnic group. Adverse impact is defined as a substantially different rate of selection in a hiring, promotion or other employment decision which works to the disadvantage of members of a race, sex or ethnic group. The EEOC has adopted the four-fifths or eighty percent rule for the determination of "substantially different" rates of selection.³⁵

32. *Castaneda*, 430 U.S. at 497 n.17, cited with approval in *Hazelwood*, 433 U.S. at 308 n.14.

33. 433 U.S. at 307-08, citing *International Bhd. of Teamsters v. United States*, 431 U.S. 324, 339 (1976).

34. Several commentators have discussed the advantages and problems with the EEOC's four-fifths rule. See, e.g., Boardman & Vining, *The Role of Probative Statistics in Employment Discrimination Cases*, 46 LAW & CONTEMP. PROBS. 189 (Autumn 1983); Meier, Sacks & Zabell, *What Happened in Hazelwood: Statistics, Employment Discrimination, and the 80% Rule*, 1984 AM. B. FOUND. RES. J. 139 (1984); Shoben, *Probing the Discriminatory Effects of Employee Selection Procedures with Disparate Impact Analysis Under Title VII*, 56 TEX. L. REV. 1 (1977); VanBowen, Jr. & Riggins, *A Technical Look at the Eighty Percent Rule as Applied to Employee Selection Procedures*, 12 U. RICH. L. REV. 647 (1978). Alternative standards for evaluating statistical evidence in Title VII cases have been proposed. See, e.g., Note, *Beyond the Prima Facie Case in Employment Discrimination Law: Statistical Proof and Rebuttal*, 89 HARV. L. REV. 387 (1975); Braun, *Statistics and the Law: Hypothesis Testing and Its Application to Title VII Cases*, 32 HASTINGS L.J. 59 (1980); Shoben, *Differential Pass-Fail Rates in Employment Testing: Statistical Proof Under Title VII*, 91 HARV. L. REV. 793 (1978); Sugrue & Fairley, *A Case of Unexamined Assumptions: The Use and Misuse of the Statistical Analysis of Castaneda/Hazelwood in Discrimination Litigation*, 24 B.C.L. REV. 925 (1983).

35. 29 C.F.R. § 1607.4D (1987).

Under this rule, a selection rate for any race, sex or ethnic group which is less than four-fifths or eighty percent of the selection rate for the group with the highest selection rate is regarded as a substantially different rate of selection.³⁶

The four-fifths rule is a useful criterion for determination of adverse impact on a particular group. The following discussion examines the application of the four-fifths rule first in general terms and then demonstrates its use in a specific case.

Consider the following simple case where the white persons, "Ws," and black persons "Bs," in Figure 1 below represent a twofold classification of the population under consideration. In Figure 1, there are three states in which the individuals can find themselves: in the general population, in the applicant pool or in the selected group. Keep in mind that the groups are not exclusively separate, thus all those in the selected group are also counted among the applicants; and all applicants are counted among the general population. Furthermore, for the purpose of applying the four-fifths rule, we may not know, nor is it always relevant, what the proportion of Bs and Ws are in the general population. The four-fifths rule is used to determine whether there has been discrimination in the passage from the applicant pool to the selected group. In particular, it is the screening device identified as hurdle A in Figure 1 that is the focus of attention. The question is: Does hurdle A have an adverse impact on black or white individuals in being selected?

FIGURE 1

General population

	Applicant pool	Selected Group
	B B W W B	
	W B B W W	/ hurdle A
	B W W B B	/ W W B W W
	W B W W B	W B W B W
	W B B B W	W W W B W
	Total	Total
	500 Bs 500 Ws	200 Bs 400 Ws

An unspecified number of Bs and Ws

36. *Id.*

The determination of adverse impact by the four-fifths rule of thumb is a four-step procedure.

- (1) Calculate the rate of selection for each group (divide the number of persons selected from a group by the number of applicants from that group).
- (2) Observe which group has the highest selection rate.
- (3) Calculate the impact ratios, by comparing the selection rate for each group with that of the highest group (divide the selection rate for a group by the selection rate for the highest group).
- (4) Observe whether the selection rate for any group is substantially less; i.e., 4/5 or 80% less than the selection rate for the highest group. If it is, adverse impact is indicated.

In relation to Figure 1 above, the first step is to identify the selection ratios for the two groups separately, i.e., to determine what the ratio of white selectees is to the number of white applicants and the ratio of black selectees to the black applicants.

$$\text{White selection ratio} = \frac{\text{White selectees}}{\text{White applicants}} = \frac{400}{500} = 80\%$$

$$\text{Black selection ratio} = \frac{\text{Black selectees}}{\text{Black applicants}} = \frac{200}{500} = 40\%$$

The next step is to observe which group has the higher selection ratio. Obviously, in this example, considering the proportion of whites to blacks in the applicant pool, whites have had better success in crossing hurdle A than blacks have had. That is, 80% of the whites passed through the screening device, but only 40% of the blacks did. Obviously, there is a difference here, but it is only with the application of the four-fifths rule that we can determine if the difference is great enough to be considered as adverse impact for the black group.

The third step is to calculate the impact ratio. This is done by taking the selection ratio for the group with the lower ratio, in this case blacks, and dividing it by the selection ratio for the group with the higher ratio. In other words, divide the adverse group ratio by the favored group ratio.

$$\text{Impact Ratio} = \frac{\text{Black selection ratio}}{\text{White selection ratio}} = \frac{40}{80} = 50\% \text{ which is less than } 4/5 \text{ or } 80\%$$

Now compare the observed impact ratio to the rule of thumb. Since the 50% which was so determined is less than the standard of 80% or 4/5, one can conclude adverse impact on the black group. That is, the

screening device adversely impacts blacks in comparison to whites and, without other evidence to the contrary, the device would be considered to be discriminatory.

It is important to note that hurdle A represents any process, device, test, procedure or practice adopted by an employer to use when an employee moves from one level of employment or training to another, including passage from outreach to intake, intake to assessment, assessment to training, training to placement, and so on.

An important feature of the four-fifths rule is its ease of use in providing a measure of adverse impact when there are more than two groups in the comparison. That is, suppose the comparison was on racial grounds; the applicant and selected groups could be divided not just on the basis of black and white selection ratios, but also on the basis of Native American, Asian and other selection rates.

The Uniform Guidelines contain two qualifications of the use of the four-fifths rule for more complex comparisons of this type. First, adverse impact determinations are to be made only on those groups which constitute 2% or more of the total labor force in the relevant labor area or 2% or more of the applicable work force.³⁷ Second, detailed comparisons for adverse impact are not required.³⁸ That is, the application of the four-fifths rule need not be made for black males, white females, black females and white males, and so on. Only comparisons based on sex alone, and race alone, need be made. Selection procedures free of adverse impact against any sex, race or ethnic group are unlikely to have an impact against a subgroup. It should be recalled that adverse impact comparisons are made with reference to the group with the *highest* selection ratio. This element becomes clearer in the case where there are more than two groups in the comparison. The following data provides an example for such a situation.

<u>Group</u>	<u>Number Applied</u>	<u>Number Selected</u>
White (Non-Hispanic)	75	45
Black	53	27
Hispanic	36	14

37. 29 C.F.R. § 1607.15A (1987).

38. *Id.*

The relevant selection ratios are given below:

$$\text{White selection ratio} = \frac{\text{White selectees}}{\text{White applicants}} = \frac{45}{75} = 0.60$$

$$\text{Black selection ratio} = \frac{\text{Black selectees}}{\text{Black applicants}} = \frac{27}{53} = 0.51$$

$$\text{Hispanic selection ratio} = \frac{\text{Hispanic selectees}}{\text{Hispanic applicants}} = \frac{14}{36} = 0.40$$

In this case the group with the highest selection ratio is the first one, whites other than Hispanic with a 60% selection rate. Thus, the determination of impact ratios is based on this group.

$$\begin{aligned} \text{Impact on Blacks} &= \frac{\text{Black selection ratio}}{\text{White selection ratio}} \\ &= \frac{51}{60} = 0.85 \text{ or } 85\% > 80\% = \text{no adverse impact} \end{aligned}$$

$$\begin{aligned} \text{Impact on Hispanics} &= \frac{\text{Hispanic selection ratio}}{\text{White selection ratio}} \\ &= \frac{40}{60} = 0.67 \text{ or } 67\% < 80\% = \text{adverse impact} \end{aligned}$$

The impact ratios shown above indicate that the selection procedure adversely impacts Hispanics but not blacks since the impact ratio for blacks is 85% which is above the 80% or 4/5 rule of thumb, and the impact ratio for Hispanics of 67% is below the 4/5 level.

Since more complex comparisons can be rather cumbersome when the calculations are written out, the information may be more compactly and efficiently handled using the tabular format. In this format the information for the total problem above might be displayed as in Table 1. In this table, the information in the top row provides the operation necessary to carry out the required calculation. Thus column (3) is determined by dividing the value of column (2) by its corresponding value in column (1), and so on. Columns (4) and (6) are informational and do not require any calculations. With experience, these columns can be left off, but they do represent essential parts of the logical process of the determination of adverse impact and should be included until some degree of familiarity with the process is achieved.

TABLE 1

Group	(1)	(2)	(3)	(4)	(5)	(6)
	Number Applied	Number Selected	Selection Ratio	Highest Ratio	Impact Ratio	Adverse Impact
Operation	Given	Given	(3)=(2)/(1)	Maximum	(5)=(3)/Max(3)	
White	75	45	60	*	—	—
Black	53	27	51	—	85	No
Hispanic	36	14	40	—	67	Yes

III

EVALUATION OF THE FOUR-FIFTHS RULE BY COMPARISON
TO A STATISTICAL CRITERION

The central question of this Article is the validity of the four-fifths rule. In particular, the four-fifths rule is compared to the theoretically sounder binomial test.³⁹ This type of statistical criterion studies

39. *Elementary Statistical Hypothesis Testing Rules.* Before a comparison of the two methods is possible we will briefly review statistical hypothesis testing principles. There are two types of errors that can be made in judging whether a company is discriminating. A Type I error means that we reject a true hypothesis. Thus, we make a Type I error when we say that the company is guilty of discrimination when in fact they are not (the differences in their hiring practice could be due to chance). A Type II error occurs when we accept a false hypothesis. This occurs if we judge the company not guilty of discrimination when in fact they are discriminating. Figure 1 sets out all the possibilities. Note that there are two correct decisions and a Type I and a Type II error.

We start with the null hypothesis, that there is no discrimination. This hypothesis must be rejected in order to show that the company is guilty of discrimination. There is always a possibility that there will appear to be discrimination when in reality none actually exists (Type I error). We must however set up some cut-off point below which we say that minority hires are so low that it seems reasonable that discrimination does exist. For example, if the applicant pool is of equal size and 100 blacks and 100 whites are hired, we are satisfied that there is no discrimination. What if 75 blacks and 100 whites are hired? Does this indicate discrimination?

Type I, Type II Errors and Correct Decisions

		Actual Situation	
		Not Guilty	Guilty
Type of Action	Employer not Discriminating	Company not discriminating and judged not guilty (correct decision)	Type I Error Company not discriminating but judged guilty
	Employer Discriminating	Type II Error Company discriminating but judged not guilty	Company discriminating and found guilty (correct decision)

Figure 1

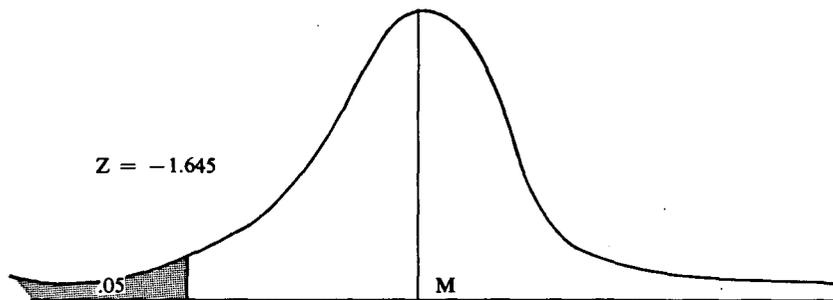
probabilities and accounts for sample size.⁴⁰ Baldus and Cole⁴¹ discuss *selection rates* which compare percentages of individuals who are selected at certain points (17% of group A vs. 80% of group B), *representation rates* (45% of pool A vs. 70% of pool B) and *actual numbers* of minorities compared to the *expected numbers*; however, the comparison of the expected numbers to the actual numbers is not made to the outcomes using the four-fifths rule which is used in most employment cases. Finkelstein⁴² also discusses the binomial criterion and uses a chi-square (χ^2) test to compare expected and actual values but he uses it in terms of jury

If we use the binomial distribution and the number of people interviewed (n) times the probability of being hired (p) is greater than or equal to 5 ($np \geq 5$), then we may use normal curve statistics to study (n) our population. Thus, if the probability of being hired is .5, the number of hires at or above which we could use the normal curve statistics is ($np = 5$) is 10. So for relatively small numbers of hires we can use the normal curve approximation to the binomial distribution.

If we know the standard deviation(s) for our data using the normal curve (figure 2) we can say that within $\pm 2s$ around the mean we include 95.5% of all possible results. Our z value

$$z = \frac{x - M}{s}$$

tells the number of standard deviations we are above or below the mean, where x = actual number of blacks hired and M = Np the expected number of hires. In the case of discrimination we are worried that the mean of black hires will be too low so we decide to put our 5% Type I error on the left side of our curve so that if the number of standard deviations that our black hires is below the white hires is less than -1.645 (this is the number of standard deviations below which we find the lowest 5% of our data—shaded area on our graph) we say that the null hypothesis thesis of no discrimination is rejected and we feel there is likely to be discrimination.



40. For very small sample sizes both the 4/5ths rule and a binomial test, based upon approximation to the normal distribution, are inadequate measures of discrimination. In the case of the 4/5ths rule, the effect of hiring or failing to hire just one person has a grossly disproportionate effect on the determination of discrimination. In the case of the binomial test here developed, statistical inferences from a normal distribution are no longer reliable. However, there are nonparametric tests (i.e., not based upon an assumption of normality in distribution) which can be used to measure discrimination. Chief among these is the chi-square (χ^2) test which has been used to measure discrimination in the context of jury selection. M. FINKELSTEIN, *QUANTITATIVE METHODS IN LAW*, 18-58 (1978). The rule of thumb is that if $npx > 5$ (number of hires times probability of minority hires, in a labor context), then normal statistics might be applicable. If $npx < 5$, then a nonparametric measure is advisable.

41. Baldus & Cole, *Quantitative Proof of Intentional Discrimination*, 1 EVALUATION QUARTERLY 53 (1977).

42. M. FINKELSTEIN, *supra* note 40.

discrimination cases and makes no mention of the four-fifths rule. The following section illustrates how to use the binomial test as the criterion for discrimination in appraising the effectiveness of the four-fifths rule.

A. Determination of Adverse Impact Using a Binomial Test

Suppose in a population of 50% white and 50% black applicants, a work force or training program consists of 135 persons, 75 white and 60 black. Using a 5% level of Type 1 error (chance of rejecting a true hypothesis—1.645 standard deviations with a one-tailed test), we can test the hypothesis that blacks and whites have equal access to the work force or the training program. In this case the test statistic is:

$$z = \frac{x - M}{s}$$

This calculated value of z is tested against the z value specified for the amount of Type I error desired. If the desired error is 1% using a one-tailed test, $z = -2.33$; if it is 5%, $z = -1.645$; and if it is 10%, $z = -1.28$; where

x = the actual number of blacks hired

$M = Np$ the expected number of blacks

$s = \sqrt{Npq}$ = standard deviation of the binomial distribution

and where

N = total number of people selected, whites plus blacks

p = the probability of selectees being black under equal access

$q = (1 - p)$ = the probability of a selectee being white under equal access

Thus in the above problem, given

$$x = 60$$

$$N = 135$$

$$p = 0.5$$

$$q = 0.5$$

consequently

$$M = 67.5$$

$$s = 5.81$$

and

$$z = \frac{60-67.5}{5.81} = -1.3$$

The conclusion here is that the hypothesis of equal access cannot be

rejected; in other words, the situation does not describe a prima facie discriminatory situation.

Now notice the results of the application of the four-fifths rule in the exact same situation as described above. Suppose that blacks and whites applied in equal numbers, and 60 blacks and 75 whites were selected. Does this constitute adverse impact using the four-fifths rule? The answer to this question is no, but the situation is right on the borderline.⁴³

$$\text{Impact ratio} = \frac{60}{75} = 80\%$$

Thus, with one less black in this example, there would be adverse impact. Notice, then, that the four-fifths rule indicates a discriminatory situation below 60 blacks, other things being constant, while standard statistical inference does not indicate any discrimination at all. In fact, one would have to drop the number of blacks to 56, when 75 whites are hired (thus a total of 131 hires) before the boundary of discrimination would appear. This is at the 5% level; if one used a 1% level, the minority hires could drop to 51 without any signs of discrimination.

$$z = \frac{56 - 65.5}{\sqrt{(0.5)(0.5)(131)}} = -1.66$$

And thus it is clear that the four-fifths rule indicates adverse impact when the number of blacks ranges from 59 down to 56 and below, while statistical inference does not register discrimination until 56 or fewer blacks are hired. Hence, in this case, the four-fifths rule signals discrimination when in fact there is none; the four-fifths rule seems to exaggerate true adverse impact.

On the other hand, consider the following situation:

With the same population of 50% black and 50% white, assume that 1600 blacks and 2000 whites are selected for the training program or work force. In this case, as that above, application of the four-fifths rule results in a finding right on the boundary;

$$\text{Impact ratio} = \frac{1600}{2000} = 80\%$$

any fewer blacks, 1599 or below, would signal adverse impact. But applying the binomial test we find:

43. In this example the impact ratios $\frac{b/B}{w/W}$ becomes b/w since $B=W$.

$$z = \frac{1600 - 1800}{\sqrt{(0.5)(0.5)(3600)}} = -6.6$$

Here -6.6 is considerably lower than our critical cut-off point of -1.65 indicating discrimination.

In contrast to the previous case, the four-fifths rule does not indicate adverse impact in a situation where there is heavy discrimination. That is, the hypothesis of equal access is rejected by a large margin using statistical inference. In fact, the number of blacks would have to increase to 1841 before there would be a result of no discrimination according to the statistical inference criterion. Thus, from 1600 to 1896, the four-fifths rule indicates no adverse impact while the binomial test indicates that there is, in fact, discrimination. The four-fifths rule, instead of exaggerating discrimination with these large numbers, is not sensitive enough to the discriminatory situation. Thus, under the four-fifths rule, Type II error is committed.

B. Boundary Conditions Using the Four-Fifths Rule and Binomial Test With 5% Type I Error

These differences will be demonstrated more generally. The following notations suggesting black-white differences will be used; however, they apply to any dichotomous division of the population:

Let:

W = the total number of whites in the population or in the applicant pool

B = the total number of blacks in the population or in the applicant pool

Thus, wherever convenient:

$r = B/W$ = the ratio of blacks to whites in the population or the applicant pool

w = the number of white participants or selectees

b = the number of black participants or selectees

z = the number of standard deviations determined to be associated with a discriminatory situation

The four-fifths rule is generalized as follows: there is adverse impact on blacks whenever

$$\frac{b/B}{w/W} < 0.8$$

Expressing the boundary condition for the four-fifths rule in terms of the number of black participants, we get

$$b = 0.8w$$

The situation for statistical inference is as follows:

$$p = \frac{B}{B + W}$$

$$q = \frac{W}{B + W}$$

$$N = w + b$$

$$x = b$$

and

$$M = pN = \frac{B}{W + B} (w + b)$$

(M is the mean calculated from the sample itself)

and

$$s = \sqrt{Npq} = [(w + b) \frac{B W}{(W + B)^2}]^{1/2}$$

(where s is the standard deviation calculated from the sample itself)

Thus, for discrimination to exist

$$\frac{x - M}{s} = \frac{b - \frac{B}{W + B} (w + b)}{[(w + b) \frac{B W}{(W + B)^2}]^{1/2}} < -Z$$

Solving this latter complex inequality for b at the boundary of significance, we get an expression in terms of x, z and r.⁴⁴

$$b = 1/2r [2w + z^2 - z \sqrt{z^2 + 4w(1 + r^{-1})}]$$

$$r = \text{ratio of blacks to whites (B/W) in the population}$$

$$r^{-1} = \text{ratio of whites to blacks (W/B) in population}$$

Table 2 presents the comparison of the four-fifths equation and the statistical inference equation (for the 5% level of error) under two possible ratios of blacks in the population: where blacks are equal to whites

44. This derivation can be found in Sobel & Ellard, *Comparison of the EEOC Four-Fifths Rule and a One, Two or Three σ Binomial Criterion* (Working Paper 80-901), Edwin L. Cox School of Business, SMU (available from the authors). The equation worked on has two roots, one at each tail of the distribution. This equation is at the negative side where black discrimination exists.

TABLE 2

Examples of Difference in Outcomes Using the Four-Fifths Rule and a Binomial Est. with 5% Type I Error and 2 Different Population Ratios ($z = -1.645$)

Ratios of Applicant Pools	Number of Majority Hires	4/5 Rule* Number of Minority Hires	Binomial Rule Number of Minority Hires
B = W	20	16	10.9
	50	40	34.9
Number of majority hires for which results are the same:	100	80	78.1
	150	120	122.83
122	200	160	168.425
B = .10W	100	8.0	4.6
	200	16.0	15.4
Number of majority hires for which results are the same: 584	500	40.0	37.9
	1000	80.0	82.9
	2000	160.0	175.78

* These figures were computed using the formulas outlined in this paper page 18.

** Line indicates crossover point, below the line the binomial criterion requires more minority hires than the 4/5 rule. Above the line the binomial criterion specifies fewer minority hires than the 4/5 rule.

and where blacks are 10% of whites. Several conclusions can be made by observing this table.

2. The error of the four-fifths rule also increases as the size of the hiring population increases. For small numbers of hires the four-fifths criterion is actually more demanding on the employer than the binomial test. For large numbers of hires the binomial test is more demanding on the employer. Thus, in comparison to the binomial test, the four-fifths rule will be *more* likely to find discrimination where it does *not* exist (Type I error) for a *small* firm, and *less* likely to find discrimination where it does exist (Type II error) for a large firm.

3. There is an interactive effect between these two sources of error. Transition points where the binomial test becomes more demanding are found at progressively higher numbers of hires as the ratio of blacks to whites gets smaller. When blacks are 50% of the population (and Type I error is set at 5% one-tailed test), the point of equality of the four-fifths rule with the binomial test is at a point (see Table 2) where there are 98 black hires and 122 white hires. After 220 hires the statistical criterion demands more blacks be hired than does the four-fifths test. When blacks are 10% of the number of whites, if we use the 5% Type I error and a one-tailed test, then the crossover point becomes 584 white hires and 47 black hires for a total of 631 hires. (Thus, if the ratio of minority groups to majority groups is small, the binomial test tends to be less demanding of minority hires than the four-fifths criterion.)

4. The four-fifths rule seems to enjoy some advantages in terms of its ease of operation for simple hand calculations. This is particularly true for handling multiple categories, e.g., black, Hispanic, and white. But, with computer application, of course, there is little concern with computational efficiency.

Although the four-fifths rule may be a convenient device for flagging discriminatory situations, it should be used with caution, particularly where population ratios are close to 50% and where sample sizes are very large.