

USE OF STATISTICAL EVIDENCE IN EMPLOYMENT DISCRIMINATION LITIGATION

AIDAN R. VINING*
DAVID C. MCPHILLIPS†
ANTHONY E. BOARDMAN†
Vancouver

This article examines the use in Canada and the United States of aggregate statistical evidence to establish and rebut allegations of employment discrimination. It reviews United States procedures, including adverse treatment and adverse impact, and discusses the use of statistical evidence in United States judicial and administrative contexts. The probity of statistical parity studies depends on two central issues: first, whether the comparison is appropriate and, second, whether the comparison is "practically significant". The article discusses problems, difficulties, pitfalls and solutions pertaining to these and related issues. It also discusses the use of regression analysis to examine wage discrimination, and evaluates the Equal Employment Opportunity Commission's "four-fifths" rule. The article then considers proof of employment discrimination in Canada. Recent Supreme Court of Canada decisions suggest Canadian human rights legislation now incorporates adverse impact discrimination. Consequently, issues relating to aggregate statistical data are now a central concern. The analytical framework presented here can facilitate correct evaluation of Canadian parity studies.

Dans cet article, les auteurs examinent la façon dont les statistiques sont utilisées au Canada et aux États-Unis pour établir ou réfuter les allégations de discrimination en ce qui concerne l'emploi. Ils étudient la façon d'opérer aux États-Unis, y compris le traitement désavantageux et l'effet désavantageux et considèrent l'utilisation des statistiques dans le contexte judiciaire et administratif. La valeur probante des études sur la parité des statistiques dépend de deux points importants: la comparaison est-elle appropriée et la comparaison est-elle importante en pratique? Les auteurs passent en revue les problèmes, difficultés, embûches et solutions qui se rapportent à ce genre de question. Ils considèrent aussi le rôle de l'analyse par régression dans l'étude de la discrimination en matière de salaires et examinent la règle des "quatre cinquièmes" de la commission sur les chances égales en matière d'emploi (Equal Employment Opportunity Commission). Les auteurs se tournent alors vers les preuves de discrimination en ce qui concerne l'emploi au Canada. Les jugements récents de la Cour suprême du Canada mènent à penser que la législation sur les droits de la personne s'applique maintenant à la discrimination de l'effet désavantageux. De ce fait, les questions

* Aidan R. Vining, of the Faculty of Business Administration, Fraser University, Burnaby, British Columbia.

† David C. McPhillips and Anthony E. Boardman, both of the Faculty of Commerce and Business Administration, University of British Columbia, Vancouver, British Columbia.

The authors are listed in reverse alphabetical order. This is not intended to reflect their relative contributions.

ayant trait aux statistiques ont pris une importance majeure. Le cadre analytique présenté par les auteurs peut permettre une évaluation plus facile des études sur la parité canadienne.

Introduction

In the last few years employment discrimination has become a topic of intense policy debate in Canada. While it has long been social policy that employment discrimination is an anathema, employers have largely been left to deal with the issue on a voluntary basis. Critics, however, are becoming increasingly concerned about the lack of progress in integrating minorities and women into all levels of the Canadian work force.¹ Broadened efforts to combat employment discrimination have recently been endorsed by the Abella Commission on Equality in Employment² and a House of Commons Special Committee Report.³

While the primary focus of media and political discussions has been on affirmative action, the calls for such a remedy presupposes effective mechanisms for establishing the existence of discrimination. Two recent Supreme Court rulings, *Re Bhinder and Canadian National Railway Company*⁴ and *Re Ontario Human Rights Commission and O'Malley and Simpsons-Sears Limited*,⁵ have considerably clarified the major issues.

¹ Much casual evidence suggests widespread, systematic, racial and sexual discrimination in Canada. However, with two possible exceptions, hard evidence on the nature and extent of employment discrimination does not exist. Most studies examine aggregate disparities in participation rates or incomes without controlling for other income-determining factors. See, for example, Report of the Commission on Equality in Employment, Judge Rosalie Abella, Commissioner, Government of Canada: Ministry of Supplies and Services (1984) (hereafter the Abella Commission), and Jac-André Boulet and Laval Lavallée, *The Changing Economic Status of Women*, Canadian Government Publishing Centre (1984). Pretty solid evidence of wage discrimination between males and females is provided by Roberta E. Robb, *Earnings Differentials Between Males and Females in Ontario, 1971* (1978), 11 *Canadian Journal of Economics* 350, although this study can be criticized for not satisfactorily controlling for different preferences. See also Walter Block and Walter Williams, *Male-Female Earnings Differentials: A Critical Appraisal* (1981), 2 *Journal of Labor Research* 385. Direct evidence of racial discrimination is found in Frances Henry and Effie Ginzberg, *Who Gets the Work? A Test of Racial Discrimination in Employment*, The Urban Alliance on Race Relations and the Social Planning Council of Metropolitan Toronto (1985). The researchers found "some form of preference for Whites took place in nearly ¼ of all job contacts . . . Whites have three job prospects to every one for Blacks". (p. 5).

² Abella Commission, *op. cit.*, footnote 1.

³ *Equality Now! Report of the Special Committee on the Participation of Visible Minorities in Canadian Society*. House of Commons, March 8th, 1984, 2nd Session, 32nd Parliament, 1983-84.

⁴ (1985), 23 D.L.R. (4th) 481, 9 C.C.E.L. 135 (S.C.C.), rev'g. (1983), 147 D.L.R. (3d) 312, 4 C.H.R.R. 284, D/1404 (Fed. C.A.), rev'g. (1981), 2 C.H.R.R. 123, D/546.

⁵ (1985), 23 D.L.R. (4th) 321, 9 C.C.E.L. 185 (S.C.C.), rev'g. (1982), 138 D.L.R. (3d) 133 (Ont. C.A.), aff'g. (1982), 133 D.L.R. (3d) 611 (Ont. Div. Ct.), aff'g. (1980), 2 C.H.R.R. 57, D/267.

Most important, in all jurisdictions a *prima facie* case of employment discrimination can now be established through the use of "adverse impact" evidence; proof of intent is not mandatory. As a consequence of these rulings, the use of statistical evidence in employment discrimination cases becomes of central importance.

The use of statistical evidence to establish a *prima facie* case of employment discrimination is well established in the United States⁶ and this has been noted in a number of recent Canadian cases. For example, in *Action Travail des Femmes and Canadian Human Rights Commission v. Canadian National*,⁷ the Canadian Human Rights Tribunal cited over thirty United States cases but only one Canadian case. The Federal Court of Appeal in *Action Travail*⁸ explicitly recognized, and indeed encouraged, the reliance on American jurisprudence in developing Canadian law. MacGuigan J. stated "[n]ot only was it not improper for the tribunal to review the wider United States experience . . . but it might have been thought to have been delinquent not to do so".⁹ However, Canadian courts and tribunals have had difficulty in applying evidentiary and statistical rules in employment discrimination cases. In particular, there remains confusion about the meaning of "adverse treatment" and "adverse impact",¹⁰ about the interpretation and efficacy of statistical data, and about the implications of the differences between the judicial and administrative approaches in the United States.

The first three parts of the article present a systematic review of the use of statistics to establish and rebut employment discrimination claims in the United States since the passage of the Civil Rights Act¹¹ of 1964. Part I deals with adverse treatment and adverse impact; Part II concerns the judicial progress in utilizing statistical data to establish discrimination, most importantly the increasing emphasis on flows and intra-company comparisons and upon multivariate analysis; and Part III examines the parallel administrative approach to dealing with these same evidentiary issues. The fourth part reviews current Canadian human rights legislation and case law on employment discrimination: specifically, establishing a *prima facie* case, rebuttal procedures, and the use of statistics.

⁶ W.B. Connolly and D.W. Peterson, *Use of Statistics in Equal Employment Opportunity Litigation* (1982); D.C. Baldus and J.W.L. Cole, *Statistical Proof of Discrimination* (1980).

⁷ (1984), 5 C.H.R.R. 396, D/2327 (Can. H.R. Tr.).

⁸ *Sub. nom. Re Canadian National Railway Co. and Canadian Human Rights Commission* (1985), 20 D.L.R. (4th) 668 (Fed. C.A.).

⁹ *Ibid.*, at p. 681.

¹⁰ Synonymous for "adverse" in this context are disparate, disproportionate, differential or disadvantageous.

¹¹ 42 U.S.C.S., ss. 2000a to 2000h-6.

This article does not discuss test validation or employment testing.¹² Other important but omitted issues are data base construction, and data reliability and validity. The article focusses on analysis of properly verified and validated data sets.

I. *Adverse Treatment and Adverse Impact in the United States*

Title VII of the United States Civil Rights Act states "[i]t shall be 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".¹³ Originally, the Act was primarily intended to address malevolent discrimination towards individuals who had been the victims of overt discrimination—normally called "adverse treatment".¹⁴ However, since *Griggs v. Duke Power Co.*¹⁵ Title VII has formed the cornerstone for dealing with discrimination against protected groups even if there appeared to be no overt intention to discriminate on the part of the employer, but where there was some aggregate evidence of discrimination—normally called "adverse impact".

In *International Brotherhood of Teamsters v. United States*,¹⁶ the Supreme Court noted:

[in an adverse treatment case] [t]he employer simply treats some people less favorably than others because of their race, color, religion, sex, or national origin . . . [adverse impact cases] involve employment practices that are facially neutral in their treatment of different groups but that in fact fall more harshly on one group than another and cannot be justified by business necessity.

Simply put, in a treatment case the employer differentiates on the basis of some illegal factor such as race or sex, while in an impact case the employer differentiates on the basis of some practice or rule that is correlated with some illegal factor such as race or sex.¹⁷ In an adverse treat-

¹² For more information on these issues see D.H. Jones, Comment (1982), 77 *Jo. American Statistical Ass.* 787; S.F. Cronshaw, *The Status of Employment Testing in Canada: A Review and Evaluation of Theory and Professional Practice* (1986), 27 *Canadian Psychology* 183.

¹³ *Supra*, footnote 11, s. 2000e-2 (a)(1)(1976).

¹⁴ Drafters of the 1964 Civil Rights Act viewed employment discrimination primarily 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". (S. REP., No. 415, 92d Congress, 1st Session. S. (1971): Senate Committee Report on Equal Employment Opportunity Enforcement Act of 1971.)

¹⁵ 401 U.S. 424 (1971).

¹⁶ 431 U.S. 324, at p. 335, n. 15 (1977).

¹⁷ Brilmayer *et al.*, *Sex Discrimination in Employer-Sponsored Insurance Plans: A Legal and Demographic Analysis* (1980), 47 *U. of Chi. L. Rev.* 505.

ment case, proof of discriminatory motive is critical while it is not required in a disparate impact case.¹⁸

We shall now outline the procedures that can be followed in adverse treatment and adverse impact cases.¹⁹ The Supreme Court has stated frequently that these procedures are not rigid or inflexible; they are frameworks or guidelines.²⁰ For example, in *Furnco Construction Corp. v. Waters*²¹ the Supreme Court noted that the procedure "was never intended to be rigid, mechanized, or ritualistic. Rather, it is merely a sensible, orderly way to evaluate the evidence in light of common experience as it bears on the critical question of discrimination".

A. Adverse Treatment

Proof of discrimination in adverse treatment cases focuses on discriminatory motivation and intent. In addition to cases of overt, malevolent intent, plaintiffs can demonstrate constructive intent. Two well-established procedures for doing so are outlined in *McDonnell Douglas Corp. v. Green*²² and *International Brotherhood of Teamsters v. United States*.²³

The procedure in an individual, covert discrimination action, which was specified in *McDonnell Douglas*, was affirmed and developed in *Furnco, Board of Trustees of Keene State College v. Sweeney*²⁴ and *Texas Department of Community Affairs v. Burdine*.²⁵ A plaintiff may establish a *prima facie* case of discrimination by showing:

- (i) that he belongs to a racial minority; (ii) that he applied and was qualified for a job for which the employer was seeking applicants; (iii) that, despite his qualifications, he was rejected, and (iv) that, after his rejection, the position remained open and the employer continued to seek applicants from persons of complainant's qualifications.²⁶

When such a constructive *prima facie* case has been established, the employer need only "articulate some legitimate, nondiscriminatory reason for the employee's rejection."²⁷ If the defendant meets this burden,

¹⁸ *International Brotherhood of Teamsters v. United States*, *supra*, footnote 16, at p. 335, n. 15.

¹⁹ C. Sullivan, M. Zimmer and R. Richards, *Federal Statutory Law of Employment Discrimination* (1980); B. Schlei and P. Grossman, *Employment Discrimination Law* (2nd ed., 1983).

²⁰ See also *International Brotherhood of Teamsters v. United States*, *supra*, footnote 16, at p. 358; *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, at p. 802, n. 13 (1973).

²¹ 438 U.S. 567, at p. 577 (1978).

²² *Supra*, footnote 20.

²³ *Supra*, footnote 16.

²⁴ 439 U.S. 24 (1978).

²⁵ 450 U.S. 248 (1981).

²⁶ *McDonnell Douglas Corp. v. Green*, *supra*, footnote 20, at p. 802.

²⁷ *Ibid.*

the plaintiff then has a final opportunity to rebut this defence by showing that the employer's stated reason was merely a pretext for discrimination.²⁸

In *Burdine*,²⁹ the Supreme Court clarified the issues relating to the burden and allocation of proof:

The ultimate burden of persuading the trier of fact that the defendant intentionally discriminated against the plaintiff remains at all times with the plaintiff . . . Establishment of the prima facie case in effect creates a presumption that the employer unlawfully discriminated against the employee . . . The burden that shifts to the defendant, therefore, is to rebut the presumption of discrimination by producing evidence that the plaintiff was rejected, or someone else was preferred, for a legitimate nondiscriminatory reason. The defendant need not persuade the court that it was actually motivated by the proffered reasons . . . It is sufficient if the defendant's evidence raises a genuine issue of fact as to whether it discriminated against the plaintiff. To accomplish this, the defendant must clearly set forth, through the introduction of admissible evidence, the reasons for the plaintiff's rejection . . . the employer need only produce admissible evidence which would allow the trier of fact rationally to conclude that the employment decision had not been motivated by discriminatory animus.

Constructive intent cases are not restricted to individual actions. Specific classes may claim that an employer engaged in a "pattern or practice" of discrimination against that protected group. Examples include *Teamsters*,³⁰ *Hazelwood School District v. United States*³¹ and, more recently, *Equal Employment Opportunity Commission v. Federal Reserve Bank of Richmond*.³² In such cases the plaintiff has "to establish by a preponderance of the evidence that . . . discrimination was the company's standard operating procedure—the regular rather than the unusual practice".³³ "Where gross statistical disparities can be shown, they alone may in a proper case constitute prima facie proof of a pattern or practice of discrimination."³⁴ The justification is that in an adverse treatment case, "[p]roof of discriminatory motive . . . can in some situations be inferred from the mere fact of differences in treatment".³⁵ In short, very large statistical disparities allow an inference of motivation and intent.³⁶ Once the plaintiff has established a *prima facie* case, "[t]he burden then shifts to the employer to defend the prima facie showing of a pattern or

²⁸ *Ibid.*, at pp. 804-805.

²⁹ *Supra*, footnote 25, at pp. 253-257.

³⁰ *Supra*, footnote 16.

³¹ 433 U.S. 299 (1977).

³² 698 F. 2d 633 (4th Cir., 1983).

³³ *Teamsters*, *supra*, footnote 16, at p. 336.

³⁴ *Hazelwood*, *supra*, footnote 31, at pp. 307-308, paraphrasing *Teamsters*, *supra*, footnote 16, at p. 339.

³⁵ *Teamsters*, *ibid.*, at p. 335, n. 15. Similarly, "[s]tatistics showing racial or ethnic imbalance are probative in a case such as this one only because such imbalance is often a telltale sign of purposeful discrimination . . ."; *ibid.*, at p. 339, n. 20.

³⁶ Of course, in such cases the personal testimony of witnesses is also relevant.

practice by demonstrating that the Government's proof is either inaccurate or insignificant".³⁷

Where a rule, test or limitation overtly discriminates against a protected group, employers may rely on the "bona fide occupational qualification" defence. The Civil Rights Act³⁸ of 1964 permits an employer to make decisions based on sex, religion or national origin "in those certain instances where [the characteristic] is a bona fide occupational qualification reasonably necessary to the normal operation of that particular business or enterprise". While the Supreme Court noted in *Dothard v. Rawlinson*³⁹ that the exception was "meant to be an extremely narrow exception to the general prohibition of discrimination on the basis of sex", the boundaries of such defences are characteristically vague.

Discrimination on the basis of religion has received special consideration. Prior to 1972, the courts treated religion in a manner similar to other protected classifications.⁴⁰ However, in 1972 Congress explicitly amended the Civil Rights Act to require an employer to make a special effort to accommodate employees' religious practices. Specifically, the amendment states an employer will be deemed to have discriminated ". . . unless [he] 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 [his] business".⁴¹

B. Adverse Impact

As a result of *Griggs*⁴² the focus of many employment discrimination cases has moved away from motivation and intent, whether overt or constructive, to the effects of an allegedly discriminatory practice. The Supreme Court held in *Griggs* that Title VII is not only concerned with an employer's "good intent or absence of discriminatory intent . . . [b]ut Congress directed the thrust of the Act to the consequences of employment practices, not simply the motivation".⁴³ The procedure to be followed in such cases was developed in *Griggs*, and refined in *Albemarle Paper Co. v. Moody*⁴⁴ and subsequent cases. Essentially, the plaintiff has to make out a *prima facie* case through presentation of statistical evidence that a facially neutral practice or rule had a disparate impact upon a

³⁷ *Teamsters, supra*, footnote 16, at p. 360.

³⁸ *Supra*, footnote 11, s. 2000e-2(e).

³⁹ 433 U.S. 321, at p. 334 (1977).

⁴⁰ See, for example, *Dewey v. Reynolds Metal Company*, 429 F.2d 324 (6th Cir., 1970); aff'd. 402 U.S. 689 (1971).

⁴¹ Equal Employment Opportunity Act of 1972, 42 U.S.C.S. s. 2000 e(j).

⁴² *Supra*, footnote 15.

⁴³ *Ibid.*, at p. 432.

⁴⁴ 422 U.S. 405 (1975).

protected group, that is, that the pattern of hiring or promotion of a protected group differed from that of the majority group. The burden of proof then shifts to the defendant who may argue that the existing pattern arose from a "job-related" criterion or that the policies or practice were justified as a "business necessity". In short, ". . . Congress has placed on the employer the burden of showing that any given requirement must have a manifest relationship to the employment in question".⁴⁵ This burden is analogous to the "legitimate, nondiscriminatory reason" defence in an adverse treatment case, although debate surrounds the issue of which is more onerous.⁴⁶

While there was no suggestion in *Griggs* that the plaintiff would have a further opportunity for rebuttal after the defendant demonstrated job-relatedness, the Supreme Court stated in *Albemarle*:⁴⁷

If an employer does then meet the burden of proving that its tests are 'job related', it remains open to the complaining party to show that other tests or selection devices, without a similarly undesirable [discriminatory] effect, would also serve the employer's legitimate interests in 'efficient and trustworthy workmanship' . . . Such a showing would be evidence that the employer was using its tests merely as a 'pretext' for discrimination.

Thus, in an impact case, the plaintiff can rebut the employer's defence by showing that the employer's reason was a pretext, again analogously to a treatment case. However, more recently, no mention of a lesser-impacting alternative was made in *New York City Transit Authority v. Beazer*.⁴⁸

C. Conclusion

While this analysis suggests a neat delineation between adverse treatment and adverse impact, a number of commentators have argued that the

⁴⁵ *Griggs*, *supra*, footnote 15, at p. 432. The Supreme Court clarified in *Albemarle* the standard of proof for establishing through validation studies whether employment tests are job related by stating ". . . discriminatory tests are impermissible unless shown, by professionally acceptable methods, to be 'predictive of 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'"; *ibid.*, at p. 431, quoting EEOC guidelines 29 CFR, s. 1607.4(c).

⁴⁶ It seems probable that the burden had been more onerous in an adverse impact case, particularly subsequent to *Dothard v. Rawlinson* in which a footnote states, "a discriminatory employment practice must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge"; *supra*, footnote 39, at p. 332, n. 14. However, the Supreme Court ruling in *New York City Transit Authority v. Beazer*, 440 U.S. 568 (1979) suggests a weakening of the "job relatedness" burden. This issue is discussed further in part II.C of this article; see also H.A. Furnish, A Path Through the Maze: Disparate Impact and Disparate Treatment Under Title VII of the Civil Rights Act of 1964 after *Beazer* and *Burdine* (1982), 23 Bost. Coll. Law Rev. 418.

⁴⁷ *Supra*, footnote 44, at p. 425, quoting *McDonnell Douglas*, *supra*, footnote 20, at pp. 804-805.

⁴⁸ *Supra*, footnote 46.

distinctions are now more illusory than real.⁴⁹ Furnish has suggested that the Supreme Court's recent opinions "represent an evolution towards an explicable merger of disparate impact and disparate treatment cases".⁵⁰ Nonetheless, some controversy remains: debate revolves around the relative evidentiary burdens and orderings of proof, and the relative rebuttals and defences in the two types of cases.⁵¹

Clearly, though, in the United States plaintiffs can make out a *prima facie* case of discrimination without necessarily establishing the employer's overt intention to discriminate. Statistical evidence alone can establish a *prima facie* case in an adverse impact case and in a class action, pattern or practice, adverse treatment case. Statistics can also be important at the rebuttal stage of an individual action, adverse treatment case. Throughout this article we shall refer to such evidence as aggregate statistical evidence. Where discriminatory practices become more covert and subtle, or where employers raise business-related defences, then aggregate statistical evidence becomes more critical. In most current cases, establishing employment discrimination would be impossible without the introduction of such evidence.

II. *Proof of Discrimination in the Judicial Context*

The United States courts have learned over the last twenty years that there are two central issues in assessing statistical evidence: (1) "what has been compared to what?" or, as it will be called here, "the parity issue", and (2) once the parity comparison is deemed to be appropriate, how does the court decide whether it is "material", that is, whether the evidence of the impact is sufficiently damaging to be "practically significant" or "legally significant".⁵² We will refer to this as the "materiality" issue. Neither question is simple, and our description of the courts' progress necessarily rationalizes and simplifies what has often been an agonizing "two steps forward, one step back" learning experience. Progress has also been retarded because the Supreme Court has often taken many years to overturn, reconcile or reinterpret the varied findings of the Courts of Appeal.

A. *The Parity Issue: From Stocks to Flows*

(1) *Introduction*

Statistics have often been presented which compare the percentage of a particular set of actual or potential protected employees with the

⁴⁹ Furnish, *loc. cit.*, footnote 46; D. Kaye, *Statistical Evidence of Discrimination* (1982), 77 *Jo. American Statistical Assoc.* 773.

⁵⁰ Furnish, *ibid.*, at p. 420.

⁵¹ Furnish, *ibid.*; J.W. Friedman, *The Burger Court and the Prima Facie Case in Employment Discrimination Litigation: A Critique* (1979), 65 *Cornell Law Rev.* 1.

⁵² This term was used in *United States v. Test*, 550 F. 2d 577, at p. 584 (10th Cir., 1976) and subsequent cases.

percentage of that group in the appropriate applicant pool or general population. Some courts, for example, have compared the percentage of black employees with the percentage of blacks in the general population (a population/work force comparison). Some other studies have compared the percentage of protected group members who are hired with the percentage of protected group members in the pool of applicants (an applicant/flow comparison). There has been much debate over which subset of protected group employees in the firm should be used and which subset of protected group members in the general population should be used for comparison. A major issue has been the circumstances under which a particular parity comparison would be probative.⁵³

Parity issues often involve an examination of the work force composition at a particular moment in time, which means that the "stock" of employees is being considered. Other studies have been concerned with the "flow" of new hirings, assignments, or promotions over time. Studies also differ in terms of whether they concern all employees in an organization or just employees at one or more particular levels of the firm. Thus the "percentage of protected group employees" may refer to the percentage in: (1) the firm's total work force, (2) the work force at a particular level or of a particular type, (3) the total new hirings, or (4) the new hirings, assignments or promotions to a particular level or particular set of jobs. The first two comparisons involve "stocks", the last two involve "flows". Where tests and other requirements have an allegedly discriminating impact, the courts typically focus on a fifth group: the percentage of protected group members in (5) the pool of potential employees, that is, those members who meet the requirement (for example, pass the test). These five groups constitute the rows in Table 1,⁵⁴ a typology of relevant parity comparisons.

Percentages of protected group employees or potential employees must be compared to the percentage of protected group members in some reference group. These reference groups form the columns of Table 1. The first three columns pertain to reference groups that are external to the firm while the last column pertains to intra-firm comparisons. Initially comparisons were made with the percentage of protected group members in the general population (column 1). In more recent cases, comparisons have been made with the percentage of protected group members in the minimum qualified work pool (column 2), the set of applicants (column 3), or to the set of total employees or set of employees at a different level

⁵³ It is important to make sure that one knows the meaning of such expressions as the "percentage of blacks hired". The phrase may mean the percentage of blacks in the set of hired individuals or, alternatively, it may mean the percentage of blacks hired from the group of black applicants.

⁵⁴ *Infra*, p. 670. While most major employment discrimination cases are presented in the table, it is not intended to be comprehensive in the sense that it includes every case that has used statistical evidence.

Table 1
A Typology of Parity Comparisons^a

Protected Group as Percentage of:	Reference Group			
	General Population	Qualified Work Force	Applicant Pool	Another Level Within Firm ^b
Employers' Total Work Force	A	F	K	P
	Teamsters ^d	Hazelwood ³¹		
	Parham ⁵⁵	Mayor of Philadelphia		
	Roman Agarwal ⁷⁹	Agarwal ⁷⁹		
Work force at a particular level	B	G	L	Q
	Teamsters ¹⁶	Agarwal ⁷⁹		Fisher ⁷⁵
	Parham ⁵⁵	Mecklenburg Parham ⁵⁵		Griggs ¹⁵
	Weber Fisher ⁷⁵	Weber		Dickerson
	Hayes ⁸⁴	Roman		Senter
	Rogers	American Nat. Bank		Teamsters ¹⁶
	Dothard ³⁹	Fed. Res. Bank of Richmond ³²		James ⁷⁶
		Williams		Roman
				Rowe ⁸⁷
				Swint Davis ⁷⁶
			Mecklenburg Chance ^{e97}	
Recent hires to whole firm	C	H	M	R
	Parham ⁵⁵	Hazelwood ³¹	Ochoa ¹¹⁰	
	Roman	Agarwal ⁷⁹	Hazelwood ³¹	
Recent hires, assignments or promotions to particular level, type or division	D	I	N	S
	Teamsters ¹⁶	Furnco ²¹	Hester ^{d64}	Rowe ⁸⁷
		Hester ^{d64}	Wade	James ⁷⁶
		Patterson	NAACP	Johnson
		Agarwal	(Ensley) ⁶⁰	Roman
		Williams	Agarwal ⁷⁹	Davis ⁷⁶
			Fed. Res. Bank of Richmond ³²	Agarwal ⁷⁹
			Amer. Nat. Bank	Dickerson
Candidates for hire or promotion	E	J	O	T
	Griggs ¹⁵		Albemarle ^{d44}	Chance ⁹⁷
	Dothard ³⁹		Green ²⁰	James ⁷⁶
	Johnson			

^a Numbers beside case names refer to footnotes in the article where citations may be found. The citations of cases in the table not otherwise referred to in the article are as follows:

Dickerson v. U.S. Steel Corp., 439 F. Supp. 55 (E.D. Pa., 1977).

Equal Employment Opportunity Commission v. American National Bank, 652 F. 2d 1176 (4th Cir., 1981).

Johnson v. Goodyear Tire and Rubber Company, 491 F. 2d 1364 (5th Cir., 1974).

Mayor of City of Philadelphia v. Educational Equality League, 415 U.S. 605 (1974).

Mecklenburg v. Montana Board of Regents of Higher Education, 13 EPD 11438 (D. of Mont., 1976).

Patterson v. American Tobacco Company, 535 F. 2d 257 (4th Cir., 1976).

Rogers v. International Paper Co., 510 F. 2d 1340 (8th Cir., 1975).

Roman v. ESB, Incorporated, 550 F. 2d 1343 (4th Cir., 1976).

Senter v. General Motors Corporation, 532 F. 2d 511 (6th Cir., 1976).

Swint v. Pullman, 539 F. 2d 77 (5th Cir., 1976).

United Steel Workers of America, AFL-CIO-CLC v. Weber, 443 U.S. 193 (1979).

Wade v. Mississippi Cooperative Extension Service, 528 F. 2d 508 (5th Cir., 1976).

Williams v. Owens-Illinois Inc., 665 F. 2d 918 (9th Cir., 1983).

^b Another level within the firm may refer to the whole operation.

^c In *Chance* comparison was made to similar positions in different organizations.

^d This type of parity study was not used but it was endorsed in this case.

or in different jobs within the same company (column 4). For example, a case would appear in cell N if a comparison is made between the percentage of recent minority hirings, assignments, or promotions to a particular level and the percentage of minority applicants. This is an example of an applicant-flow comparison. Of the twenty cells in Table 1, three (K, L and P) are necessarily empty. We now discuss other cells, considering each column in turn.

(2) *General Population Comparisons*

Cells A and B concern population/work force comparisons. Cell A involves comparison between all employees and the general population, while cell B involves comparison between employees at a particular level and the general population. *Parham v. Southwestern Bell Telephone Co.*⁵⁵ provides a good example of a cell A comparison. Evidence showed that while fewer than two per cent of the company's employees were non-white, approximately twenty-two per cent of the population of Arkansas was black.⁵⁶ The United States Supreme Court endorsed the use of such comparisons in *Teamsters*.⁵⁷

Statistics showing racial or ethnic imbalance are probative in a case such as this one only because such imbalance is often a telltale sign of purposeful discrimination; absent explanation, it is ordinarily to be expected that nondiscriminatory hiring practices will in time result in a work force more or less representative of the racial and ethnic composition of the population in the community from which employees are hired.

While population/work force comparisons have often been used to establish a *prima facie* case of discrimination, the courts have become increasingly aware of many significant, potential problems. Many of these problems are overcome by the use of alternative probative statistics. However, one problem that has never been completely resolved is the

⁵⁵ 433 F. 2d 421 (8th Cir., 1970).

⁵⁶ *Ibid.*, at pp. 424, 426 n. 4.

⁵⁷ *Supra*, footnote 16, at p. 339, n. 20.

appropriate definition of "community from which employees are hired" or "relevant geographic region".

One problem that has eventually been resolved is the "pre-Act/post-Act" problem. Population/work force comparisons are unable to differentiate between discriminatory behaviour prior to the effective date of Title VII of the Civil Rights Act of 1964 and subsequent employment practices. As the Supreme Court noted in *Teamsters*, once a plaintiff has established a *prima facie* case "[a]n employer might show . . . that the claimed discriminatory pattern is a product of pre-Act hiring rather than unlawful post-Act discrimination . . .".⁵⁸ To deal with this problem, statistics are required which consider the "flow" of new hirings since the effective date of the Act (row 3, Table 1).⁵⁹ Such data, however, require time to accumulate. While the effective date of the Act was July 2, 1965 for most employees, Title VII was not made applicable to some employees until much later. For example, it was not made applicable to the Personnel Board of Jefferson County and the government agencies it serves until March 24, 1972.⁶⁰ The United States experience suggests that "pre-Act/post-Act" considerations remained important for a considerable number of years after the passage of the initial legislation, and is only now beginning to fade, more than twenty years after the Civil Rights Act was enacted.⁶¹

Another important case involving general population comparisons is *Griggs*. The employer required new employees to possess a high school diploma and to register satisfactory scores on standardized tests in order to qualify for placement in any department except labour. Statistical evidence was presented which showed that in North Carolina "while 34% of white males had completed high school, only 12% of Negro males had done so".⁶² Additional evidence showed that the pass rates on the standardized tests was lower for blacks than for whites.⁶³ There was, in effect, a substantial disparity between the percentage of black males in the pool of potential employees and the percentage of black males in the general population.

⁵⁸ *Ibid.*, at p. 360.

⁵⁹ In *Parham*, *supra*, footnote 55, at p. 424, the court ruled that pre-Act data may "serve only as a basis for comparison".

⁶⁰ See *N.A.A.C.P., Ensley Branch v. Seibels*, 13 EPD 11504 (N.D. Ala., 1977).

⁶¹ See A.H. Beller, *The Impact of Equal Opportunity Policy on Sex Differentials in Earnings and Occupations* (1982), 94 American Economic Association Papers and Proceedings 171; F. Welch, *Affirmative Action and Its Enforcement* (1981), 93 American Economic Association Papers and Proceedings 127; N.S. Barrett, *Obstacles to Economic Parity for Women* (1982), 94 American Economic Association Papers and Proceedings 160.

⁶² *Supra*, footnote 15, at p. 430, n. 6.

⁶³ *Ibid.*

(3) *Qualified Work Force/Relevant Labour Market Comparisons*

The courts have increasingly recognized that comparison should be made with the qualified work force in the general population (column 2) rather than with the general population *per se* (column 1) when such qualifications are essential for the job. For example, in *Hester v. Southern Railway Company*⁶⁴ the Court of Appeal ruled:

. . . [c]omparison with general population statistics is of questionable value when we are considering positions for which, as here, the general population is not presumptively qualified . . . A more significant comparison might perhaps be between the percentage of blacks in the population consisting of those able to type 60 wpm or better and the percentage hired into the Data Typist position by Southern.

The Supreme Court reaffirmed this view in *Hazelwood*.⁶⁵

In *Teamsters*, the comparison between the percentage of Negroes on the employer's work force and the percentage in the general areawide population was highly probative, because the job skill there involved—the ability to drive a truck—is one that many persons possess or can fairly readily acquire. When special qualifications are required to fill particular jobs, comparisons to the general population (rather than to the smaller group of individuals who possess the necessary qualifications) may have little probative value.

The Supreme Court ruled that “a proper comparison was between the racial composition of Hazelwood's teaching staff and the racial composition of the qualified public school teacher population in the relevant labor market”.⁶⁶ Although the court did not actually delineate the relevant labour market, it noted that the definition was important and enunciated some factors that should be considered.⁶⁷

(4) *Applicant Pool Comparisons*

A different type of comparison is involved when the court uses as a reference group the percentage of protected group applicants. Since it makes little sense to compare the work force “stock” with the composition of the applicant pool, cells K and L are empty. Evidence focuses on a comparison of the “flow” of employees, whether recent hirings, promotions, or assignments, with the applicant pool. These applicant/flow cases are found in cells M and N.

⁶⁴ 497 F. 2d 1374, at p. 1379, n. 6 (5th Cir., 1974).

⁶⁵ *Supra*, footnote 31, at p. 308, n. 13.

⁶⁶ *Ibid.*, at p. 308. See also at p. 310: “The record in this case showed that for the 1972-1973 school year, Hazelwood hired 282 new teachers, 10 of whom (3.5%) were Negroes; for the following school year it hired 123 new teachers, 5 of whom (4.1%) were Negroes. Over the two-year period, Negroes constituted a total of 15 of the 405 new teachers hired (3.7%).” These percentages could have been compared with the percentage of teachers in St. Louis and the city of St. Louis who were black (15.4%) or the percentage of teachers in the St. Louis County alone who were black (5.7%).

⁶⁷ *Ibid.*, at p. 311, 311 n. 17, 312.

Although applicant/flow data were not available in *Hazelwood*, the court realized that such data would "be very relevant".⁶⁸ Several other cases have suggested that applicant/flow data are more useful than work force/population comparisons.⁶⁹ For example, in *Ochoa v. Monsanto Company*⁷⁰ there was a substantial disparity between the percentage of Mexican-Americans employed and the percentage of Mexican-Americans in the general population, but this evidence was found not to constitute *prima facie* discrimination, partly because the company's applicant/flow data showed that Mexican-Americans had recently been hired at a faster rate than applicants of any other racial group.⁷¹

The Supreme Court ruled in *Albemarle* that where job requirements or tests are at issue, the plaintiff can make out "a prima facie case of discrimination . . . [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".⁷² Thus, while *Griggs* compared the percentage of blacks who were selected for hire with the percentage of blacks in the general population (cell E), *Albemarle* compared the percentage of blacks selected for hire and the percentage of black applicants (cell O).

One important problem with applicant/flow data is that employers may "chill" the applicant pool. This problem is described in detail in *Teamsters*:⁷³

If an employer should announce his policy of discrimination by a sign reading 'Whites Only' on the hiring-office door, his victims would not be limited to the few who ignored the sign and subjected themselves to personal rebuffs. The same message can be communicated to potential applicants more subtly but just as clearly by an employer's actual practices—by his consistent discriminatory treatment of actual applicants, by the manner in which he publicizes vacancies, his recruitment techniques, his responses to casual or tentative inquiries, and even by the racial or ethnic composition of that part of his work force from which he has discriminatorily excluded members of minority groups.

An exclusive reliance by the courts on applicant/flow data would provide a tremendous incentive for employers to manipulate the applicant pool in ways that would be detrimental to protected groups. In order to obtain a high selection rate for a protected group, a company would have an incentive to discourage minority applications, contrary to the intent of a reasonable antidiscriminatory law.

⁶⁸ *Ibid.*, at p. 308, n. 13.

⁶⁹ The appropriate use of applicant/flow data is also suggested by *Hester* which suggested the plaintiffs could compare "the percentage of blacks among those applying for a particular position and the percentage of blacks among those hired for the position"; *supra*, footnote 64, at p. 1379.

⁷⁰ 335 F. Supp. 53 (S.D. Tex., 1971), *aff'd*, 473 F.2d 318 (5th Cir., 1973).

⁷¹ *Ibid.*, at p. 59 (F. Supp.).

⁷² *Supra*, footnote 44, at p. 425.

⁷³ *Supra*, footnote 16, at p. 365.

(5) *Intra-Firm Comparisons*

The fourth column of Table 1 compares the percentage of protected group members employed or potentially employed at one level with the percentage of protected employees at some other hierarchical level of the firm or in the entire firm. Typically, the purpose for making a claim in this category is to show that the employer engaged in a general pattern and practice of discriminating against minority employees in making promotions. Cell Q comparisons are often regarded as probative. In *Teamsters*, for example, the evidence concerned company-wide discrimination:⁷⁴

. . . the company had 6,472 employees. Of these, 314 (5%) were Negroes and 257 (4%) were Spanish-surnamed Americans. Of the 1,828 line drivers, however, there were only 8 (0.4%) Negroes and 5 (0.3%) Spanish-surnamed persons, and all of the Negroes had been hired after the litigation had commenced . . . A great majority of the Negroes (83%) and Spanish-surnamed Americans (78%) who did work for the company held the lower paying city operations and serviceman jobs, whereas only 39% of the nonminority employees held jobs in those categories.

In the last sentence of this passage, the proportion of protected employees at a less desirable position is implicitly compared with the proportion of protected employees at a more desirable level.

More recently in *Fisher v. Procter & Gamble Manufacturing Co.*⁷⁵ the plaintiffs argued that in making cell B and cell Q comparisons they had established a *prima facie* discrimination case, while the company countered by arguing that a cell G comparison was most relevant. The Appeal Court held that the company would be estopped from using the qualified work force argument in rebutting disparity claims when the firm itself engaged in substantial training for these skilled positions:⁷⁶

When a company adopts a policy and practice of hiring in at low level unskilled jobs and promoting to upper-level positions based upon training received and skills developed at the plant itself, it cannot convincingly challenge the *prima facie* showing under the *Hazelwood* 'qualifications' dicta. Where skills are commensurate with company training, we will approve statistical comparisons between racial make-up in key positions and racial composition in the total work force.

(6) *Conclusion*

Comparisons to the general population suffer from three flaws. First, they ignore the skill level of the general population; second, the "relevant geographic region" may be difficult to define; and, third, some of the data may be based on pre-Act rather than post-Act hiring practices. However, progress has been made on all three problems. The first is well

⁷⁴ *Ibid.*, at pp. 337-338.

⁷⁵ 613 F. 2d 527 (5th Cir., 1980), *cert. denied*, 449 U.S. 1115 (1981).

⁷⁶ *Ibid.*, at p. 544. Similar arguments are presented in *James v. Stockham Valves & Fittings Co.*, 559 F. 2d 310, at p. 341 (5th Cir., 1977), *cert. denied*, 434 U.S. 1034 (1978) and *Davis v. Califano*, 613 F. 2d 957, at p. 965 (D.C. Cir., 1979).

recognized by the courts and, when qualifications are important for the job, comparison is now made to one of the other reference groups. The "pre-Act/post-Act" problem is also well recognized: for the most part, pre-Act data are relevant only as a basis for comparison.⁷⁷ It was often necessary, therefore, to examine new hirings (flows) rather than the existing work force (stocks), although such data took years to develop in some cases. Comparisons to the applicant pool are generally regarded as better than comparisons to the qualified work force.⁷⁸ However, a serious potential problem is that employers may discourage protected group applicants and thereby "chill" the applicant pool. More recently, many cases have made comparison to all or part of the employer's total work force (cells Q and S).

No single comparison is without its faults. Perhaps the most important conclusion that emerges from an examination of the table is that few cases rely on a single type of comparison. *Teamsters*, for example, is listed in four cells. The most frequently cited case, *Agarwal v. McKee & Co.*⁷⁹ appears in eight cells. In the United States many plaintiffs do not rely upon just a few statistical comparisons to shift the burden of proof; a complete picture of the employer's practices can be developed.

B. *The Materiality Issue: From Exclusion to Regression*

(1) *Introduction*

The success of a plaintiff in using statistical data to establish a prima facie case depends crucially on the magnitude of the disparity—it must be "material". Table 2⁸⁰ presents the five major criteria or tests that have been applied to determine whether the disparity is sufficient to shift the burden to the defendant. The tests become more sophisticated as one moves down the column and, generally, this is the sequence the courts have followed over time. The progression appears to be a function of both the courts' increasing statistical sophistication and the changing nature of the employment discrimination cases that have come before the courts.

(2) *Rule of Exclusion*

When initially concerned with employment discrimination cases, the courts found consideration of discrimination in jury selection (equal protection) cases provided a valuable perspective. In several jury cases, no protected group member had ever had been in the selected jury group. These "rule of exclusion" cases date back over one hundred years to

⁷⁷ See *supra*, footnote 59.

⁷⁸ For the reasons behind this view see Mr. Justice White's minority opinion in *Dothard*, *supra*, footnote 39, at pp. 347-348, which was also applicable to *Hazelwood*, *supra*, footnote 31.

⁷⁹ 16 EPD 8301 (N.D. Calif., 1977), aff'd. 644 F. 2d 803 (9th Cir., 1981).

⁸⁰ *Infra*, p. 677.

Table 2
A Typology of Materiality Criteria^a

Criterion	Cases			
Rule of Exclusion	James ⁷⁶ Watkins Wade Teamsters ¹⁶	Pierre Franks Bing	Valentino ¹²⁴ American Nat. Bank	
Overwhelming or Gross Disparity	Parham ⁵⁵ Teamsters ¹⁶ Dothard ³⁹ Fisher ⁷⁵	Hazelwood ³¹ Senter James ⁷⁶ Johnson	Rowe ⁸⁷ Swint Griggs ¹⁵ Dickerson	Kaplan ⁸⁷ Hayes ⁸⁴ Watkins Ochoa ¹¹⁰
Absolute Disparity Standard	Dothard ³⁹ Test ⁵² Davis ⁷⁶ Chance ⁹⁷	Sims Whitus		
Univariate Significance Tests	NAACP (Ensley) ⁶⁰ Hazelwood ³¹ Albemarle ⁴⁴ Roman	Agarwal ⁷⁹ Dickerson Swint Contreras	Fed. Res. Bank of Richmond ³²	
Multivariate Statistical Tests	James ⁷⁶ Wade Mecklenburg Agarwal ⁷⁹	Stastny ¹¹⁶ Presseisen Dickerson Vuyanich ¹²³	Greenspan ¹²² NAACP (Ensley) ⁶⁰ Valentino ¹²⁴ Keely ¹²⁴	Fed. Res. Bank of Richmond ³² Wilkins ¹²⁴

^a Numbers beside case names refer to footnotes in the article where citations may be found. The citations of cases in the table not otherwise referred to in the article are as follows:

- Bing v. Roadway Express, Inc.*, 485 F. 2d 441 (5th Cir., 1973).
Contreras v. City of Los Angeles, 656 F. 2d 1667 (9th Cir., 1983).
Dickerson v. U.S. Steel Corporation, 439 F. Supp. 55 (E.D. Pa., 1977).
Franks v. Bowman Transportation Company, 495 F. 2d 398 (5th Cir., 1974).
Johnson v. Goodyear Tire and Rubber Company, 491 F. 2d 1364 (5th Cir., 1974).
Mecklenburg v. Montana Board of Regents of Higher Education, 13 EPD 11438 (D. of Mont., 1976).
Pierre v. Louisiana, 306 U.S. 354 (1938).
Presseisen v. Swarthmore College, 442 F. Supp. 593 (E.D. Pa., 1977).
Roman v. ESB, Incorporated, 550 F. 2d 1343 (4th Cir., 1976).
Sims v. Georgia, 389 U.S. 404 (1967).
Swint v. Pullman, 539 F. 2d 77 (5th Cir., 1976).
Wade v. Mississippi Cooperative Extension Service, 528 F. 2d 508 (5th Cir., 1976).
Watkins v. Scott Paper Company, 530 F. 2d 1159 (5th Cir., 1976).
Whitus v. State of Georgia, 385 U.S. 545 (1967).

Neal v. Delaware,⁸¹ where over seventeen per cent of the population was black but "no colored citizen had ever been summoned as a juror in the courts of the State".

⁸¹ 103 U.S. 370, at p. 397 (1880). Similar exclusionary evidence was presented in *Hill v. Texas*, 316 U.S. 400 (1942); *Patton v. Mississippi*, 332 U.S. 463 (1947); *Hernandez v. Texas*, 347 U.S. 475 (1954); *Arnold v. North Carolina*, 376 U.S. 773 (1964).

Total exclusion of a protected group has also been observed in employment discrimination cases. In *Teamsters*, for example, the Supreme Court observed that "the company's inability to rebut the inference of discrimination came not from a misuse of statistics but from 'the inexorable zero'".⁸² When a protected group has been totally excluded, discrimination is conspicuous and it obviates the need for detailed, narrow attacks.

(3) *Overwhelming Disparity*

More frequently in early discrimination cases, protected groups were not totally excluded but were "overwhelmingly" underrepresented.⁸³ An early example is *United States v. Hayes International Corporation*⁸⁴ in which evidence showed that while only six (0.65 per cent) of the 924 office and technical employees were black and fourteen (4.6 per cent) of the 299 new hirings were black, thirty per cent of the general population was black. The court ruled "[t]hese lopsided ratios are not conclusive proof of past or present discriminatory hiring practices; however, they do represent a *prima facie* case".⁸⁵ In *Teamsters*, the Supreme Court said that "[e]vidence of longlasting and gross disparity between the composition of a work force and that of the general population thus may be [legally] significant . . .".⁸⁶ Disparities reported in *Griggs* and *Hazelwood* were considered to be "substantial". Other synonyms have included "marked", "compelling", "wholly disproportionate", and "stark".⁸⁷

In cases where there was "total exclusion" or an "overwhelming" disparity, a cursory examination of the employer's practices suggested manifest discrimination. It was entirely reasonable for the Supreme Court to take the position, as it did in *Castaneda v. Partida*,⁸⁸ a jury selection case, that "[i]f a disparity is sufficiently large, then it is unlikely that it is due solely to chance or accident, and, in the absence of evidence to the

⁸² *Supra*, footnote 16, at p. 342, n. 23. Data showed "[w]ith one exception. . .the company and its predecessors *did not employ a Negro on a regular basis as a line driver until 1969*. And, as the Government showed, even in 1971 there were terminals in areas of substantial Negro population where all of the company's line drivers were white"; *ibid.*, at p. 337.

⁸³ *Ibid.*, at p. 342, n. 23.

⁸⁴ 456 F. 2d 112, at p. 120 (5th Cir., 1972).

⁸⁵ *Ibid.* (Emphasis added).

⁸⁶ *Supra*, footnote 16, at p. 339, n. 20. (Emphasis added). The Supreme Court also used the term "gross" in *Dothard*, *supra*, footnote 39, at p. 330, n. 12.

⁸⁷ "Marked": see, e.g., *Griggs*, *supra*, footnote 15, at p. 429; *Rowe v. General Motors Corporation*, 457 F. 2d 348, at p. 357, n. 18 (5th Cir., 1972); "compelling": *Fisher*, *supra*, footnote 75, at p. 544; "wholly disproportionate": *Kaplan v. Intern. Alliance of Theatrical, Etc.*, 525 F. 2d 1354, at p. 1358 (9th Cir., 1975); "stark": *Village of Arlington Heights v. Metropolitan Housing Development Corporation*, 429 U.S. 252, at p. 266 (1977).

⁸⁸ 430 U.S. 482, at p. 494, n. 13 (1977).

contrary, one must conclude that racial . . . factors entered into the selection process". Soon thereafter, in *Teamsters* and *Hazelwood*, the Supreme Court allowed "gross" statistical disparities to establish *prima facie* cases in "pattern or practice" employment discrimination cases. While statistical disparities do not conclusively prove employer intent to discriminate it may allow an inference of intent when the disparity is sufficiently large or "overwhelming". Precedent for such an inference dates back to *Yick Wo v. Hopkins*,⁸⁹ in which evidence showed that seventy-nine out of eighty non-Chinese applications for laundry permits were successful while about 200 Chinese nationals were not.

(4) *Absolute Disparity Standard*

As long as the disparities were very large there was no statistical problem. When discrimination evidence was less than "overwhelming," however, the courts encountered considerable difficulty in establishing appropriate criteria and in using more complex statistical analyses. In practice, the courts often resorted to the use of *ad hoc* comparisons of percentage or ratio differences. We have grouped these cases in row three of Table 2 under the heading "absolute disparity standard".⁹⁰ In *Swain v. Alabama*,⁹¹ another jury selection case, "[t]he evidence was that while Negro males over 21 constitute 26% of all males in the county in this age group, only 10 to 15% of the grand and petit jury panels drawn from the jury box since 1953 have been Negroes . . .". The court admitted that "[w]e cannot say that purposeful discrimination based on race alone is satisfactorily proved by showing that an identifiable group in a community is underrepresented by as much as 10%".⁹² Of course, the Supreme Court should have said "underrepresented by as much as 16%". In any case, the ten to sixteen per cent range of disparity became a standard for comparison in future decisions.⁹³

⁸⁹ 118 U.S. 356 (1886).

⁹⁰ See Table 2, *supra*, p. 677. For a discussion of problems with this criteria in the context of jury selections see David Kairys. Joseph Kandane and John Lehoczky. *Jury Representativeness: A Mandate for Multiple Source Lists* (1977), 65 Cal. Law Rev. 776.

⁹¹ 380 U.S. 202, at p. 205 (1965).

⁹² *Ibid.*, at pp. 208-209.

⁹³ For example, the Tenth Circuit in *United States v. Test*, *supra*, footnote 52, at p. 587, stated:

In the present cases the maximum disparity demonstrated by defendants between the percentages of blacks and Chicanos in the voting-age community and on the master jury rolls was approximately 4%. Since this figure is well below the 10-16% range of disparity approved in *Swain*, the district court properly concluded defendants had failed to establish a *prima facie* case of systematic exclusion and accepted the government's general explanations and asseverations of good faith in rebuttal.

Absolute disparity comparisons have been used in employment discrimination cases attacking allegedly discriminatory requirements.⁹⁴ In *Griggs*⁹⁵ twelve per cent of black males had completed high school while thirty-four per cent of white males had done so. This particular arbitrary percentage difference established a *prima facie* case and became a benchmark in subsequent employment discrimination actions.⁹⁶

Sometimes a court takes an additional step and computes the selection ratio which is then compared to some arbitrary number. The selection ratio of whites to the protected group ranged from 1.3 in *Chance v. Board of Examiners*⁹⁷ to 3.5 in *Bridgeport Guardians, Inc. v. Members of Bridgeport Civil Service Commission*.⁹⁸ The courts may also compute failure rates. For example, in *Davis v. Washington*⁹⁹ the failure rate was fifty-seven per cent for blacks and thirteen per cent for whites so that "[t]he disparity disclosed in this case—more than four to one—is larger than differences held sufficiently disproportionate in other cases".¹⁰⁰ Considerable care should be exercised with selection and failure rates: the outcome depends to a large extent upon whether one computes the pass ratio or the fail ratio, and on the magnitude of the statistics.¹⁰¹

In time, these mechanistic and arbitrary approaches were deemed to be inadequate to handle increasingly complex cases. As cases involving smaller disparities came before the courts, the use of the concept of statistical significance became essential.

(5) *Univariate Significance Tests*

The concept of statistical significance was introduced in the 1974 Equal Employment Opportunity Commission Guidelines¹⁰² and a five per cent level of significance was employed in *Albemarle*.¹⁰³ To say that a relationship is statistically significant at, for example, the five per cent

⁹⁴ The four-fifths per cent rule, which will be critiqued in Part III, falls in this category.

⁹⁵ *Supra*, footnote 15, at p. 430, n. 6.

⁹⁶ For example, *Dothard, supra*, footnote 39, at p. 330, n. 12.

⁹⁷ 458 F. 2d 1167 (2d Cir., 1972).

⁹⁸ 482 F. 2d 1333 (2d Cir., 1973).

⁹⁹ 512 F. 2d 956 (D.C. Cir., 1975), rev'd. 426 U.S. 229 (1976).

¹⁰⁰ *Ibid.*, at p. 960 (F. 2d).

¹⁰¹ For example, in *Davis, ibid.*, the failure ratio was four to one but the pass ratio was 1.66 to one, a major difference. If the failure rate for blacks had been 4% and the failure rate for whites had been 1%, the ratio would still have been four to one, but there would have been much less reason for concluding that blacks had been systematically excluded.

¹⁰² Uniform Guidelines on Employee Selection Procedures (1978), 29 C.F.R. s. 1607 (1982).

¹⁰³ *Supra*, footnote 44, at pp. 430, 431, 432, 437.

level of significance means that there is a five per cent or lower probability that the observed result, or one more disparate, was due to chance.¹⁰⁴ An alternative but equivalent approach is to calculate the number of standard deviations¹⁰⁵ by which the observed result differs from the result expected had there been no discrimination. Since the five per cent level of significance is equivalent to 1.96 standard deviations, it is consistent with the "two or three standard deviations" criterion approved by the Supreme Court in *Castaneda*.¹⁰⁶

The importance of statistical significance was firmly established in *Hazelwood*¹⁰⁷ and is now well accepted. Nonetheless, the courts have recognised that there are problems. Statistical significance tests are not entirely satisfactory in some circumstances because of the effect of sample size. In large samples, small percentage differences will be statistically significant. Knowledge of whether or not a finding is statistically significant does not necessarily imply importance, relevance or practical significance. A conclusion that a relationship is statistically significant may say little of consequence about its magnitude and importance when the sample size is very large.¹⁰⁸

When samples are very small, large differentials are necessary to obtain statistically significant results. Furthermore, the test is not as pow-

¹⁰⁴ Classical hypothesis testing requires specification of a null hypothesis (that there is no relationship) and an alternative hypothesis (that there is a relationship). The analyst also specifies a level of significance of, say, 5% which is the (maximum) probability of rejecting the null hypothesis when, in fact, the null hypothesis is true, *i.e.* concluding there is a relationship when in fact there is none. In other words, the existence of a relationship that is significant at the 5% level of significance means if there was no discrimination, the actual outcome (or one more disparate) would occur by chance no more than five times out of a hundred or one time in twenty, on average.

The influence of a factor may be measured by the t-statistic or F statistic. If the sample size is large and the t statistic exceeds 1.96 the null hypothesis is rejected at the 5 percent level of significance for a two-sided alternative. Alternatively, we could say the null hypothesis is rejected at the 95 per cent level of confidence.

¹⁰⁵ The standard deviation is a measure of dispersion. For a normally distributed random variable there is 68.26% probability a variable is within one standard deviation of the mean, 95.44% probability of being within two standard deviations and a 99.74% probability of being within three standard deviations. Alternatively, one can say there is less than 5% probability (4.56%) of being more than 2 standard deviations from the mean, and less than 1% probability of being more than 3 standard deviations from the mean.

¹⁰⁶ *Supra*, footnote 88, at p. 496, n. 17.

¹⁰⁷ *Supra*, footnote 31.

¹⁰⁸ It is more informative to present the *p*-value and this is evidenced in many recent cases. An essentially equivalent approach involves calculating the prediction interval for different *p*-values and examining whether the observed result falls outside the interval and, if so, by how much; see Kaye, *loc. cit.*, footnote 49. The *p*-value or prob-value, is the probability of obtaining the sample result or one more damaging to the null hypothesis given the null hypothesis is true, *i.e.* it equals the probability of obtaining the observed disparity or worse given there was, in fact, no discrimination. The null hypothesis of no discrimination is rejected if the prob-value is less than the specified level of significance, frequently set at the 5% level.

erful as one based on larger samples: there may be a high probability of concluding that there is no discrimination when, in fact, it exists. (Statisticians refer to this as a Type II error.) The courts have been wary of using data in such situations. For example, in *Teamsters*¹⁰⁹ the Supreme Court suggested that an employer could defeat a prima facie showing of discrimination by demonstrating that "during the period it is alleged to have pursued a discriminatory policy it made too few employment decisions to justify the inference that it had engaged in a regular practice of discrimination". However, the courts have been inconsistent in establishing what constitutes "small".¹¹⁰

Two other problems concern the design and nature of the data collection. First, the data being analysed may represent the entire population, for example, the data may cover all employees in an organization. In such instances it may be difficult to justify classical statistical inference.¹¹¹ Second, an employer's selection process is unlikely to be entirely random because employers will generally hire and promote the most qualified candidates. A parity study may omit (productivity) variables that might explain a detected disparity, and may thereby be discredited. As the court noted in *United States v. Ironworkers Local 86*,¹¹² the use of statistics must be conditioned on the "absence of variables which would undermine the reasonableness of the inference of discrimination which is drawn". It may be necessary to consider variables such as experience, job-related qualifications, seniority, plant hiring needs and whether employees choose to work in that particular plant or job. If any one of these variables is incorporated into a parity study in addition to race or another protected group variable, the analysis, by definition, is multivariate, and multivariate statistical methodologies must be used.

(6) *Multivariate Statistical Tests*

Univariate statistical methodologies allow one to examine the effect

¹⁰⁹ *Supra*, footnote 16, at p. 360.

¹¹⁰ In *Ochoa v. Monsanto Company*, 473 F. 2d 318 (5th Cir., 1973) during the relevant period there were 684 applicants, of which 11 were Mexican-American. Fifty-six people were hired, of which one was Mexican-American. Here, Mexican-Americans were hired at a faster rate than other applicants, but the crucial issue is that "the smallness of the numbers demonstrates that the Court was not compelled to allow such statistical showing to set in train the usual presumptions or to make a finding of preference thereon"; (at p. 320). In contrast, a work force of approximately 500 was regarded as "sizable" in *Fisher*, *supra*, footnote 75, at p. 544 (F. 2d), and 870 was large in *Castaneda*, *supra*, footnote 88, at p. 496.

¹¹¹ In classical statistical inference, a sample is analysed and inferences are made about the population from which the sample was taken. Some employment discrimination data sets may contain data on all individuals in a firm, and consist, therefore, of the entire population of employees. In this case, statistical inference is generally inappropriate and a data analytic approach should be taken. Statistical inference is appropriate, however, if the data can be viewed as being generated by a process in which employment decision policies are applied to random inputs or if the data can be regarded as a random sample from a population that spans a longer time period.

¹¹² 443 F. 2d 544, at p. 551 (9th Cir., 1971).

of only one factor (whether an individual belongs to a protected group or not) on hiring, promotion or selection practices. Multivariate methodologies allow for the inclusion of a number of other relevant factors. Thus one can examine the effect of being in a protected group, after controlling for other explanatory variables.¹¹³

Multivariate analysis has been used primarily in cases dealing with wage discrimination. Usually the question before the court has been whether the average salary of protected group members is significantly lower than the average salary of the majority group, after controlling for productivity factors such as experience and education. The most common technique used to address this issue has been estimation of salary equations by multiple regression analysis.¹¹⁴

The use of multiple regression to establish a *prima facie* case of employment discrimination has raised a number of statistical issues. First, variables may have been measured incorrectly; for example, productivity

¹¹³ While multivariate analysis of selection and promotion procedures are natural extensions of the analyses described above, this approach has generally not been followed. For more information on this methodology see Baldus and Cole, *op. cit.*, footnote 6, ss. 8.021 and 8.3.

¹¹⁴ A standard general model is:

$$\text{Salary} = f(\text{group status, productivity, group status} \times \text{productivity}). \quad (1)$$

If G is a binary variable equal to unity for the protected group and zero for the majority group, and P is a productivity measure, a linear regression model can be estimated in which salary (S) is the dependent variable:

$$S = B_0 + B_1G + B_2P + B_3(GP) + e. \quad (2)$$

Group status variables (G) include race, sex, or any other protected category. Productivity variables (P) may include experience, qualifications, education, and other on-the-job measures such as ratings and absenteeism. The variable "group status \times productivity" (GP) indicates that some of the discriminatory group status variables, such as race and sex, interact with some of the productivity variables. It allows for the possibility that productivity has different effects on the salaries of protected and nonprotected group members.

Discrimination is evidenced if the coefficients of the group status variables indicate a significantly lower mean salary for the protected group than for the majority group. If $G = 1$ for protected group members and $G = 0$ for majority group members, discrimination is evidenced when B_1 is negative and significant. It is also evidenced if the coefficients of the interaction terms indicate that productivity factors have a significantly smaller effect on the protected group's salary than on the majority group's salary, when, for example, B_3 is negative and significant.

Based on equation (2), it is useful to compute what a protected group member would have received if he or she had been paid according to the same schedule as a majority group member with similar characteristics. The difference between this predicted salary and the actual salary is a measure of salary inequality adjusted for productivity characteristics. For an application of this methodology see J.R. Antos and S. Rosen, *Discrimination in the Market for Public School Teachers* (1975), 3 *Jo. Econometrics* 123, at pp. 147-148.

For an alternative, unconventional but potentially useful approach see D.A. Conway and H. V. Roberts, *Reverse Regression, Fairness, and Employment Discrimination* (1983), 1 *Jo. of Business and Economic Statistics* 75.

is difficult to measure accurately.¹¹⁵ Second, some of the explanatory variables, including productivity, may themselves be functions of discrimination.¹¹⁶ Third, important explanatory variables may be omitted from the regression (for example, productivity, seniority, merit). Fourth, the mathematical functional form may be incorrect.¹¹⁷ In addition, small sample size and "pre-Act/post-Act" problems may exist, and there may be problems associated with the regression, such as non-normality of the error terms, heteroscedasticity and multicollinearity.¹¹⁸ All of these issues have been reviewed in detail elsewhere.¹¹⁹

Many courts encountered considerable difficulty deciding the complex statistical issues, and cases often pitted one set of expert witnesses against another. In 1979 Federal Judge Winner complained that "[t]he courts have unintentionally opened a Pandora's Box by using the word 'statistics' instead of 'percentages,' because now Title VII cases [have become] contests between college professor statisticians who revel in discoursing about advanced statistical theory".¹²⁰ He further noted that

¹¹⁵ Techniques are available to overcome this problem if multiple indicators of productivity are available; see K.G. Jöreskog, A General Method for Estimating a Linear Structural Equation System, in A.S. Goldberger and O.D. Duncan (eds.), *Structural Equation Models in the Social Sciences* (1973), 85; A.E. Boardman, B.S. Hui and H. Wold, The Partial Least Squares-Fix Point Method of Estimating Interdependent Systems with Latent Variables (1981), A10 *Communications in Statistics* 613.

¹¹⁶ Merit variables such as skill level and merit rating are subjective and are often tainted; see, for example *James v. Stockham Valves and Fittings Co.*, *supra*, footnote 76, at p. 332; *Stasny v. Southern Bell Tel. & Tel. Co.*, 458 F. Supp 314, at p. 323, n. 3 (W.D. N.C., 1978).

¹¹⁷ For example, the interaction term (GP) in equation (2) in footnote 114, *supra*, may be incorrectly omitted. Another type of specification error occurs when a linear model is estimated, like equation (2), but the true model is linear in logarithms (log-linear), exponential or has some other non-linear functional form. For more discussion of the latter point see G.P. McCabe, Jr., *The Interpretation of Regression Analysis Results in Sex and Race Discrimination Problems* (1980), 34:4 *The American Statistician* 212.

¹¹⁸ For an explanation of these problems see an econometrics text such as J. Johnston, *Econometric Methods* (2nd ed., 1972).

¹¹⁹ See Connolly and Peterson, *op. cit.*, footnote 6; Baldus and Cole, *op. cit.*, footnote 6; M.O. Finkelstein, *The Judicial Reception of Multiple Regression Studies in Race and Sex Discrimination Cases* (1980), 80 *Col. L. Rev.* 737; A.B. Smith and T.G. Abram, *Quantitative Analysis and Proof of Employment Discrimination*, [1981] *U. Ill. L. Rev.* 33, at p. 34, n. 3 (for a complete bibliography of such studies to that date); F.M. Fisher, *Multiple Regression in Legal Proceedings* (1980), 80 *Col. L. Rev.* 702; D.W. Peterson, *Pitfalls in the Use of Regression Analysis for the Measurement of Equal Employment Opportunity* (1981), 5 *Int. Jo. Policy Analysis and Information Systems* 43; A. Barnett, *An Underestimated Threat to Multiple Regression Analyses Used in Job Discrimination Cases* (1982), 5 *Ind. Rel. L. J.* 156; D.E. Bloom and M.R. Killingsworth, *Pay Discrimination Research and Litigation: The Use of Regression* (1982), 21 *Industrial Relations* 318; A.E. Boardman and A.R. Vining, *The Role of Probative Statistics in Employment Discrimination Cases*, (1983), 46:4 *Law and Cont. Problems* 189.

¹²⁰ *Otero v. Mesa County Valley School District No. 51*, 470 F. Supp. 326, at p. 331 (D. Col., 1979).

“[j]udges are quite handicapped in trying to understand this testimony”.¹²¹ Recent cases, however, indicate that some courts have become very sophisticated. In *Greenspan v. Automobile Club of Michigan*¹²² the judge was able to reconcile the plaintiff’s and the defendant’s apparently contradictory regression findings. In *Vuyanich v. Republic National Bank of Dallas*,¹²³ Judge Higginbotham wrote a 165-page opinion that addressed many key statistical issues and serves as a useful model. And in a number of cases the courts discredited the plaintiffs’ regression analyses for omitting potentially important variables.¹²⁴

(7) Conclusion

When evidence of total exclusion or “overwhelming” underrepresentation was presented, discrimination was easily established, and sophisticated statistical methodologies were unnecessary. As discrimination became less overt, the courts sought refuge in arbitrary threshold values. As recently as 1975 it was argued that in the United States that “[t]he crux of the current method of proof in discrimination cases is the presentation of percentage differences sufficiently substantial to suggest bias”.¹²⁵ But, in time, this changed. The rulings in such cases as *Albemarle*,¹²⁶ *Castaneda*¹²⁷ and *Hazelwood*¹²⁸ showed that the Supreme Court began to accept the importance of statistical significance. Recent cases have involved more sophisticated methodologies as emphasis has shifted away from parity studies to wage studies. While regression analysis has sometimes been used incorrectly and has created problems for the courts, considerable progress has been made. Over time, the courts have become increasingly sophisticated and more familiar with statistical methodologies and the technical issues.¹²⁹

¹²¹ *Ibid.*, at p. 331, n. 2.

¹²² 22 FEP Cases 184 (E.D. Mich., 1980).

¹²³ 24 EPD 31480 (N.D. Tex., 1980).

¹²⁴ *Keely v. Westinghouse Electric Corp.*, 404 F. Supp. 573 (E.D. Mo., 1975); *Agarwal v. McKee & Co.*, *supra*, footnote 79; *Wilkins v. University of Houston*, 654 F. 2d 388 (5th Cir., 1981); *Valentino v. United States Postal Service*, 674 F. 2d 56 (D.C. Cir., 1982).

¹²⁵ Note, *Beyond the Prima Facie Case in Employment Discrimination Law: Statistical Proof and Rebuttal* (1975), 89 Harv. L. Rev. 387, at p. 393.

¹²⁶ *Supra*, footnote 44.

¹²⁷ *Supra*, footnote 88.

¹²⁸ *Supra*, footnote 31.

¹²⁹ S.E. Fienberg, *The Increasing Sophistication of Statistical Assessments as Evidence in Discrimination Litigation* (1982), 77 Jo. American Statistical Assoc. 784, described the court’s progress as “remarkable” (at p. 784), but Kaye *loc. cit.*, footnote 49, is quite critical.

C. *Rebuttal*

Once the plaintiff has established a *prima facie* case of treatment or impact discrimination, the defendant may show that the plaintiff has not proven discrimination because the plaintiff's evidence is incorrect, misleading or inappropriate.¹³⁰ For example, an employer may show that while a comparison of the composition of his work force with the qualified work pool indicates adverse impact on blacks, post-Act applicant flow data show no adverse impact. Much of the preceding discussion includes reference to the circumstances under which an employer can or cannot rebut a plaintiff's *prima facie* case by using this type of defence.

Alternatively, the defendant may acknowledge that his employment practices have an adverse impact but then attempt to justify his employment practices on the grounds of business-relatedness. *Dothard* well illustrates the set of issues in an impact case. Evidence showed:

. . . [the height] requirement would operate to exclude 33.29% of the women in the United States between the ages of 18-79, while excluding only 1.28% of the men between the same ages. The 120-pound weight restriction would exclude 22.29% of the women and 2.35% of the men in this group. When the height and weight restrictions are combined, Alabama's statutory standards would exclude 41.13% of the female population while excluding less than 1% of the male population.¹³¹

On the basis of such evidence the plaintiff established a *prima facie* case of unlawful sex discrimination and the burden shifted to the defence which argued "[these] requirements . . . have a relationship to strength, a sufficient but unspecified amount of which is essential to effective job performance as a correctional counsellor".¹³² This business necessity defence convinced neither the District Court nor the Supreme Court because "the appellants produced no evidence correlating the height and weight requirements with the requisite amount of strength thought essential to good job performance".¹³³ Further, no evidence was produced of a lesser-impacting alternative:

If the job-related quality that the appellants identify is bona fide, their purpose could be achieved by adopting and validating a test for applicants that measures strength directly. Such a test, fairly administered would fully satisfy the standards of Title VII because it would be one that 'measure[s] the person for the job and not the person in the abstract'.¹³⁴

No such evidence was produced and the Supreme Court ruled that the defendant had violated Title VII.

¹³⁰ *Teamsters*, *supra*, footnote 16, at p. 360, uses the descriptors: "inaccurate or insignificant".

¹³¹ *Supra*, footnote 39, at pp. 329-330.

¹³² *Ibid.*, at p. 331.

¹³³ *Ibid.*

¹³⁴ *Ibid.*, at p. 332, quoting *Griggs*, *supra*, footnote 15, at p. 436.

The Supreme Court's more recent decision in *Beazer*¹³⁵ highlights how easily the defendant can demonstrate "job relatedness" in some situations. The plaintiffs challenged the validity of a New York Transit Authority anti-narcotics rule that excluded all methadone users from all jobs in the Transit Authority. Evidence showed that some Transit jobs were not safety-critical, and that most methadone users who had been in a methadone program for one year were free from drug use. However, in a one-sentence conclusion the court held that the application of the anti-narcotics rule to exclude *all* methadone users from *all* jobs was "job related".

In *Dothard* the defendant challenged a regulation that explicitly excluded women from "contact positions", which applied to approximately seventy-five per cent of the correctional counsellors jobs. The employer relied on the "*bona fide* occupational qualification" defence. While the Supreme Court noted that this exception was meant to be extremely narrow, it concluded that the regulation fell within the narrow ambit of the "*bona fide* occupational qualification" exception on the grounds that there was "substantial testimony from experts on both sides of this litigation that the use of women as guards in 'contact' positions under the existing conditions in Alabama maximum-security male penitentiaries would pose a substantial security problem, directly linked to the sex of the prison guard".¹³⁶

D. Conclusion

The courts have become increasingly sophisticated in their analysis of statistical data. At the same time, however, they have been reluctant to rely exclusively on statistical evidence. In *Teamsters*¹³⁷ the Supreme Court noted the usefulness of statistics "depends on all of the surrounding facts and circumstances". Drawing a conclusion of "legal significance" entails an integration of statistical and non-quantitative information; consequently, some statisticians have argued that judges should learn and understand the results of a proper Bayesian analysis.¹³⁸ While there are currently valid pragmatic objections to a Bayesian approach, its adoption would hold considerable promise.

¹³⁵ *Supra*, footnote 46.

¹³⁶ *Supra*, footnote 39, at p. 336.

¹³⁷ *Supra*, footnote 16, at p. 340. The Fifth Circuit has warned that the statistical significance criterion "must not be interpreted or applied so rigidly as to cease functioning as a guide and become an absolute mandate or proscription"; *United States v. Georgia Power Company*, 474 F. 2d 906, at p. 915 (5th Cir., 1973). Similarly, in *Test*, the Tenth Circuit argued that "[t]he mathematical conclusion that the disparity between these two figures is 'statistically significant' does not, however, require an *a priori* finding that these deviations are 'legally significant' . . ."; *supra*, footnote 52, at p. 584.

¹³⁸ Classical statistics makes inferences and performs hypothesis tests on the basis of the sample information alone. In contrast, a Bayesian approach combines the sample data

The development of employment discrimination law has taken a considerable percentage of appeal courts' (including the Supreme Court's) time and energy. In 1981, for example, 13,750 civil rights cases were tried in the federal courts.¹³⁹ The resource costs of such a judicial strategy are substantial. Partially to offset these costs Congress undoubtedly had hoped that it could rely on the Equal Employment Opportunity Commission to play a prominent role in enforcing employment discrimination laws. Indeed, the budget of the Commission grew from three million dollars in 1966 to one hundred and forty-two million in 1981.¹⁴⁰ However, as the following part suggests, some attempts by the Commission and other agencies to manage administratively the problem of establishing discrimination have resulted in numerous criticisms.

III. *Proof of Discrimination in Administrative Contexts: The Equal Employment Opportunity Commission and "The Four-Fifths Rule"*

In *Griggs*,¹⁴¹ the Supreme Court held that Equal Employment Opportunity Commission administrative guidelines on the measurement of employment disparity should be viewed with "great deference". In 1978 the various administrative agencies adopted the four-fifths or "80% rule" whereby a selection rate or promotion rate of a protected group less than four-fifths of the selection or promotion rate of any other group will generally be regarded as evidence of adverse impact.¹⁴² On the other hand an employer who meets the four-fifths rule will not normally be prosecuted under Title VII.

According to Blumrosen,¹⁴³ "*Griggs* was the stick, threatening an employer with a finding of illegality, an injunction, and financial conse-

with information obtained prior to the empirical observations. Prior distributions are revised by the sample data to produce posterior distributions. In this way, objective or subjective prior beliefs about the extent of employment discrimination are formally included in the analysis. See Fienberg, *loc. cit.*, footnote 129; Kaye, *loc. cit.*, footnote 49; Morris H. de Groot, *Optimal Statistical Decisions* (1970).

¹³⁹ R. Follett and F. Welch, *Testing for Discrimination in Employment Practices* (1983), 46:4 *Law and Cont. Problems* 171.

¹⁴⁰ *Ibid.*

¹⁴¹ *Supra*, footnote 15, at p. 434.

¹⁴² Specifically the 80% rule states:

A selection rate for any race, sex, or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact.

(Uniform Guidelines on Employee Selection Procedures (1978), 29 C.F.R. s. 1607.4D (1982)).

¹⁴³ A.W. Blumrosen, *The Bottom Line Concept in Equal Employment Opportunity Law* (1980), 12 *N. Carolina Cent. L. Jo.* 1, at p. 6.

quences if his business practices perpetuate the inferior status of minorities and women. The 'bottom line' [80% rule] is the carrot, rewarding the employer whose practices 'mirror' the congressional purpose'. Certainly, the specific nature of the criterion encourages employers to follow the four-fifths rule and thus reduce the probability of liability. It should be stressed, however, that while the rule has been interpreted as an affirmative action remedy, our purpose here is to evaluate it as a mechanism for making a *prima facie* determination of the existence of employment discrimination.

There are several problems that seriously weaken the four-fifths rule's usefulness. First, it is comparatively arbitrary. Why four-fifths rather than seventy-five per cent or ninety per cent? Next, one can consider it in light of the previous section. We note the four-fifths rule measures applicant-flow and, therefore, applies to cells M or N of Table 1, while the four-fifths criterion itself belongs to the "arbitrary" percentage difference category in Table 2.¹⁴⁴ The problems associated with applicant-flow statistics and "arbitrary" percentage differences discussed above pertain directly to the four-fifths rule.

More importantly, the rule ignores the concepts of chance and statistical significance. It can be shown that the four-fifths rule and the statistical significance criterion indicate discrimination in quite different situations.¹⁴⁵ When sample sizes are small, there is a high probability that an employer will be found to be discriminating under the four-fifths rule, when in fact, he is not discriminating (a Type I error); and, when sample sizes are large, there is a high probability that an employer will be held harmless due to compliance with the four-fifths rule when, in fact, he is discriminating against a group of employees (a Type II error).¹⁴⁶

Possibly being aware of the statistical problems associated with strict adherence to the four-fifths rule, the promulgating agencies attempted to clarify their employee selection guidelines in the form of a series of questions and answers.¹⁴⁷ Of particular interest are those situations in which unusually large or small numbers of people are selected. With regard to small number cases, the document states:¹⁴⁸

¹⁴⁴ Table 1, *supra*, p. 670; Table 2, *supra*, p. 677.

¹⁴⁵ When few individuals are selected the probability that the protected group might claim adverse impact is much higher under the four-fifths rule than under the statistical significance rules. When about 200 people are selected, the rules are identical. For larger selections the statistical significance rule is more stringent for the employer than is the four-fifths rule. See Boardman and Vining, *loc. cit.*; footnote 119, at p. 216.

¹⁴⁶ For calculation of these probabilities see A.E. Boardman, Another Analysis of the EEOC 'Four-Fifths' Rule (1979), 25 Management Science 770.

¹⁴⁷ 44 Fed. Reg. 11,996 (March, 1979) (Adoption of Questions and Answers to Clarify and Provide a Common Interpretation of the Uniform Guidelines on Employee Selection Procedures).

¹⁴⁸ *Ibid.*, at p. 11,999, Q. 21.

... Generally, it is inappropriate to require validity evidence or to take enforcement action where the number of persons and the difference in selection rates are so small that the selection of one different person for one job would shift the result from adverse impact against one group to a situation in which that group has a higher selection rate than the other group.

In practice, this modification of the four-fifths rule makes little difference. With respect to large sample cases, the interpretations state that "relatively small differences in selection rates may nevertheless constitute adverse impact if they are both statistically and practically significant".¹⁴⁹ Since the "clarifications" fail to define practical significance, they do not solve the evidentiary question.

An additional problem with the four-fifths rule is that it ignores everything besides the selection rate: the courts have long recognized that the simple parity studies presented in Table 1¹⁵⁰ are appropriate only when no other factor is relevant. It should be clear from the above discussion of regression analysis that the courts and employers recognize that skill and productivity are important and legitimate factors in hiring and promotion decisions. In contrast, the guidelines ignore these variables. Furthermore, the four-fifths rule ignores wage discrimination: consequently, a corporation could pay majority group workers more than protected group workers whether majority group workers are more or less productive than, or equally productive as, protected group workers.

The rule may also encourage undesirable corporate responses. It has already been demonstrated that there will be incentives to "chill" the applicant pool. Furthermore, since a corporation has a strong incentive to comply with the four-fifths rule in order to avoid potentially large legal and administrative costs, it may not hire the most productive personnel, and the average work force quality may decline. Conversely, some firms may have desired to discriminate in the past but were discouraged from doing so because of ignorance relating to the decision rules used by the courts or the administrative agencies. The four-fifths rule could be pushed to its limit, but there is a risk with this strategy: if a company does in fact discriminate, an injured individual or group may still bring a Title VII case without involving the Equal Employment Opportunity Commission.

The problem is more complex with regard to some junior positions. Most large organizations attempt to create a pool of junior employees from which future promotions will be made. While certain skills and qualifications may be unnecessary for the available position, they may be important for subsequent promotions. "Warm-body" comparisons almost completely ignore this possibility. This type of rule, therefore, is appropriate only where it is unlikely that individuals in junior positions within the organization will be promoted to higher positions. Since these "warm-

¹⁴⁹ *Ibid.*, Q. 22.

¹⁵⁰ *Supra*, p. 670.

body" rules discourage the hiring of workers with such potential, they would eventually encourage organizations to recruit for more skilled or senior positions from outside the organization.

Some commentators, most notably the Abella Commission,¹⁵¹ have proposed that Canada should follow the United States in terms of setting up an administrative type of agency like the Equal Employment Opportunity Commission. While our analysis in and of itself does not demonstrate the superiority of a judicial approach over an administrative approach, the four-fifths rule is symptomatic of many of the problems associated with a bureaucratic, administrative approach to establishing employment discrimination.

IV. *Proof of Employment Discrimination in Canada*

In Canada, issues relating to employment discrimination come under provincial or federal Human Rights legislation,¹⁵² depending on the constitutional responsibility for the industry in which the employer operates.¹⁵³ All ten provincial legislatures and the Federal Parliament have enacted statutes which broadly prohibit employment discrimination but which differ in terms of procedure¹⁵⁴ and substance. Of critical importance,

¹⁵¹ *Supra*, footnote 1.

¹⁵² The first serious Canadian discrimination legislation was enacted in Ontario in 1944. I.A. Hunter, *Human Rights Legislation in Canada: Its Origin, Development and Interpretation* (1976), 15 U.W.O. L. Rev. 21, at p. 25, emphasizes the importance of this Act:

The 1944 Racial Discrimination Act was an important pioneering statute because, for the first time, a legislature had explicitly declared that racial and religious discrimination was against public policy, and thenceforth the judiciary could not simply subordinate human rights to commerce, contract or property.

While this milestone Act pre-dated United States civil rights legislation by many years, this and subsequent Acts, such as Equal Pay Acts, have been severely limited in their applications.

Current human rights legislation in Canada can be traced from the Ontario Human Rights Act of 1962/1964. Other jurisdictions have basically adopted that model during the intervening period. Despite the 1944 "headstart" on the United States it is only during the last ten years that courts, tribunals and commissions have begun to put some teeth into the statutes.

¹⁵³ Under the Constitution Act, 1867, as amended, employees in banks, the post office, airlines, inter-provincial railroads, telecommunications, federal crown corporations and federal public servants come under the jurisdiction of federal legislation. On the other hand, provincial legislation regulates employees of municipalities, educational institutions, natural resource companies (*e.g.*, forestry, mining), provincial crown corporations and provincial public servants. Roughly 90% of Canadians are covered by provincial laws.

¹⁵⁴ The various acts are administered by administrative bodies (Human Rights Commissions in most jurisdictions) with enforcement of the Act usually covering four phases: complaint, investigation, settlement and adjudication. B. Wilson, *Law in Society: The Principle of Sexual Equality* (1983), 13 Man. L. J. 221, at pp. 229-230, has described the process thus:

statutes might differ on the questions of whether intent is required, which groups are protected (for example, sexual orientation) and whether there is an applicable saving clause (for example, “*bona fide* occupational requirement”).

A. *Establishing a Prima Facie Case—the Relevance of Intent*

Canadian tribunals and courts have not directly adopted the United States terminology of adverse treatment and adverse impact. In the Canadian context the following terms have been substituted for adverse treatment evidence: “direct”, “intentional”; and for adverse impact: “indirect”, “unintentional”, “institutional”, “structural”, “systemic”, and “adverse effect”.¹⁵⁵ To avoid ambiguity we will employ the United States terminology.

While it is clear that adverse treatment is covered in all the statutes, the applicability of adverse impact has varied over time and among jurisdictions. The crucial issues before the courts have been whether adverse treatment alone or *both* adverse treatment and adverse impact is applicable, and what defences are acceptable. Additionally, several shifts in direction within some jurisdictions contributed to the complicated web of law. However, as a result of the Supreme Court’s decisions in *Bhinder*¹⁵⁶ and *O’Malley*¹⁵⁷ these issues have been clarified considerably.

Prior to these decisions, the major issue had been whether the legislation in question required proof of intent. Some statutes were interpreted as encompassing adverse impact and not necessarily requiring proof of intent. For example, the Ontario legislature explicitly included adverse impact through an amendment to the Code in 1982:¹⁵⁸

10. A right of a person under Part I is infringed where a requirement, qualification or consideration is imposed that is not discrimination on a prohibited ground

Once an acceptable complaint is filed, the Commission, not the complainant, controls the process. The Commission must investigate the complaint and attempt to settle it. If attempts at settlement fail, then the Commission may (or may not) ask the Minister to appoint a Board of Inquiry to look into the complaint. The Minister may (or may not) do so.

Access to Board of Inquiry or the courts is controlled by the Commissions and, in some jurisdictions, by the Minister.

For a discussion on whether complainants have civil causes of action for discrimination see Ian A. Hunter, *Civil Actions for Discrimination* (1977), 55 Can. Bar Rev. 106.

¹⁵⁵ See Wilson, *ibid.*: *Bhinder*, *supra*, footnote 4; *O’Malley*, *supra*, footnote 5; Abella Commission, *op. cit.*, footnote 1; J. Gandz and J.C. Rush, Human Rights and the Right Way to Hire (1983), 48 Business Quarterly 70; H.C. Jain, Race and Sex Discrimination in Employment in Canada (1982), 37(2) Relations Industrielles 344.

¹⁵⁶ *Supra*, footnote 4.

¹⁵⁷ *Supra*, footnote 5.

¹⁵⁸ Ontario Human Rights Code, R.S.O. 1980, c. 340, as amended in 1981, S.O. 1981, c. 53, proclaimed June 15, 1982.

but that would result in the exclusion, qualification or preference of a group of persons who are identified by a prohibited ground of discrimination and of whom the person is a member, except where,

- (a) the requirement, qualification or consideration is a reasonable and *bona fide* one in the circumstances; or
- (b) it is declared in this Act that to discriminate because of such ground is not an infringement of a right.

Provisions in Manitoba and, prior to 1984, in British Columbia were also widely interpreted as allowing the introduction of adverse impact evidence to establish a *prima facie* case of discrimination.¹⁵⁹

However, other statutes were the subject of much legal debate. The Ontario provision in effect before the 1982 amendment and the existing Federal legislation (which were the focus of *O'Malley* and *Bhinder*, respectively) provide examples. The Ontario provision read as follows:¹⁶⁰

- 4. (1) No person shall,
 - (a) refuse to refer or to recruit any person for employment;
 - (b) dismiss or refuse to employ or to continue to employ any person; . . .
 - (g) discriminate against any employee with regard to any term or condition of employment, because of race, creed, colour, age, sex, marital status, nationality, ancestry or place of origin of such person or employee.

The Canadian Human Rights Act¹⁶¹ contains two related provisions:

- 7. It is a discriminatory practice, directly or indirectly,
 - (a) to refuse to employ or continue to employ any individual, or
 - (b) in the course of employment, to differentiate adversely in relation to an employee,
 on a prohibited ground of discrimination.
- 10. It is a discriminatory practice for an employer or an employee organization
 - (a) to establish or pursue a policy or practice, or
 - (b) to enter into an agreement affecting recruitment, referral, hiring, promotion, training, apprenticeship, transfer or any other matter relating to employment or prospective employment,
 that deprives or tends to deprive an individual or a class of individuals of any employment opportunities on a prohibited ground of discrimination.

¹⁵⁹ *British Columbia Human Rights Commission v. College of Physicians and Surgeons* (1976), B.C. Board of Inquiry; *Foster v. B.C. Forest Products* (1979), B.C. Board of Inquiry, upheld on appeal [1979] 2 W.W.R. 289 (B.C.S.C.); *Grafe v. Sechelt Building Supplies* (1979), B.C. Board of Inquiry; *Holloway v. Clairco Foods Limited* (1983), B.C. Board of Inquiry; *Dhaliwal and Director, Human Rights Code v. B.C. Timber Ltd.* (1983), 4 C.H.R.R. D/1520 (B.C. Board of Inquiry); *Froese v. Pine Creek School Div.* 30 (1976), Manitoba Board of Adjudication; *McDonald v. Knit-Rite Mills Ltd.* (1984), 5 C.H.R.R. D/1949 (Manitoba Board of Adjudication); *Osborne v. Inco Ltd.*, [1984] 5 W.W.R. 228 (Man. Q.B.).

¹⁶⁰ *Supra*, footnote 158.

¹⁶¹ S.C. 1976-77, c. 33.

A number of tribunals or courts have held that these or similar provisions allowed adverse impact evidence.¹⁶² Conversely, other decisions interpreted the statutes as being limited to adverse treatment.¹⁶³

In both *Bhinder* and *O'Malley*, the Supreme Court of Canada unanimously held that such language covered adverse impact. In *O'Malley*,¹⁶⁴ McIntyre J., writing for the entire court, observed:

To take the narrower view and hold that intent is a required element of discrimination under the Code would seem to me to place a virtually insuperable barrier in the way of a complainant seeking a remedy. It would be extremely difficult in most circumstances to prove motive, and motive would be easy to cloak in the formation of rules which, though imposing equal standards, could create, as in *Griggs v. Duke Power Co.* (1970), 401 U.S. 424, injustice and discrimination by the equal treatment of those who are unequal: *Dennis v. U.S.* (1949), 339 U.S. 162 at p. 184.

In *Bhinder*,¹⁶⁵ he reiterated these conclusions. Dickson C.J.C., who dissented on other grounds, also adopted this view.¹⁶⁶

Of particular importance are the court's comments in *O'Malley*¹⁶⁷ concerning the meaning and scope of human rights legislation throughout Canada:

Legislation of this type is of a special nature, not quite constitutional but certainly more than the ordinary—and it is for the courts to seek out its purpose and give it effect. The Code aims at the removal of discrimination. This is to state the obvious. Its main approach, however, is not to punish the discriminator, but rather to provide relief for the victims of discrimination. It is the result or the effect of the action complained of which is significant. If it does, in fact, cause discrimination, if its effect is to impose on one person or group of persons obligations, penalties, or restrictive conditions not imposed on other members of the community, it is discriminatory.

¹⁶² *Singh v. S.I.S.* (1977), Ontario Board of Inquiry; *Adler v. Ontario Board of Police Commissioners* (1978), Ontario Board of Inquiry; *Colfer v. Ottawa Board of Commissioners of Police* (1979), Ontario Board of Inquiry; *Re Rocca Group Ltd. and Muise* (1979), 102 D.L.R. (3d) 529 (P.E.I.S.C.); *Malik v. Ministry of Government Services* (1981), 2 C.H.R.R. D/374 (Ontario Board of Inquiry); *Bhinder v. C.N.R.* (1981), *supra*, footnote 4 (Human Rights Tribunal); *Rand and Canadian Union of Industrial Employees and Sealy Eastern Ltd. Upholstery Division* (1981), 3 C.H.R.R. D/938 (Ontario Board of Inquiry); *Carson v. Air Canada* (1982), 3 C.H.H.R. D/818 (Fed. Hum. R. Trib.); *Marcotte v. Rio Algom* (1982), 3 C.H.R.R. D/988 (Fed. Hum. R. Trib.); *Ontario Human Rights Commission et al. v. Borough of Etobicoke*, [1982] S.C.R. 202; *Christie v. Central Alberta Dairy Pool* (1985), 6 C.H.R.R. D/2488 (Alberta Board of Inquiry).

¹⁶³ *Bhinder*, *supra*, footnote 4 (F.C.A.); *O'Malley*, *supra*, footnote 5 (Ont. C.A.).

¹⁶⁴ *Supra*, footnote 5, at pp. 331 (D.L.R.), 198 (C.C.E.L.).

¹⁶⁵ *Supra*, footnote 4, at pp. 501 (D.L.R.), 146 (C.C.E.L.).

¹⁶⁶ He stated: "I concur with Justice McIntyre . . . in concluding that the definitions of discriminatory practices in the *Canadian Human Rights Act*, ss. 7 and 10, extend to both unintentional and adverse effect discrimination." *ibid.*, at pp. 484 (D.L.R.), 149-150 (C.C.E.L.).

¹⁶⁷ *Supra*, footnote 5, at pp. 329 (D.L.R.), 196 (C.C.E.L.).

These decisions suggest that in the absence of a clear, unambiguous legislative requirement of proof of intent, human rights legislation throughout Canada must be currently interpreted as covering adverse impact discrimination as well as adverse treatment. Consequently, the statistical issues relating to parity and materiality, which were examined earlier in the United States context, are now central issues in Canadian courts and tribunals.

B. *Rebuttal of a Prima Facie Case*

Having settled the impact issue with dispatch, the Supreme Court in *Bhinder* and *O'Malley* focussed the debate on rebuttal evidence. First, the court adopted the approach suggested in such decisions as *Marcotte*¹⁶⁸ and *Canada Safeway v. Manitoba Food and Commercial Workers Union and Manitoba Human Rights Commission*¹⁶⁹ in holding that a finding of adverse impact only establishes a *prima facie*, but rebuttable, case of discrimination: an employer may justify his employment practices on the grounds of business necessity.¹⁷⁰ In *Marcotte* the majority of the Review Tribunal held that, although the company's (facially neutral) practice of providing subsidized housing only to employees in certain job classifications had an adverse impact upon women, it was justified by business necessity.¹⁷¹ In *Canada Safeway*, Wright J. of the Manitoba Court of Queen's Bench held that although a "no beards" policy had an adverse impact on men, it could be justified on the basis of reasonable occupational requirement.

Second, the Supreme Court set down different rebuttal burdens to be met by the defendant employer based on whether the relevant legislation contains a "saving" clause that explicitly permits a "*bona fide* occupational qualification" rebuttal. Where an applicable saving clause is included in the legislation, *Bhinder* held that the employer's business policy must meet the two-part test set out by the Supreme Court in the adverse treatment case, *Ontario Human Rights Commission v. Etobicoke*.¹⁷² The subjective element of the test requires the employer to demonstrate that he acted honestly, in good faith and with a sincerely held belief that the

¹⁶⁸ *Supra*, footnote 162.

¹⁶⁹ (1984), 9 D.L.R. (4th) 330, 5 C.H.R.R. D/2133 (Man. Q.B.); aff'd. (1984), 13 D.L.R. (4th) 314, 6 C.H.R.R. D/2831 (Man. C.A.).

¹⁷⁰ The Ontario Court of Appeal in *O'Malley*, *supra*, footnote 5, confused the issue of treatment versus impact discrimination with the issue of a *bona fide* occupational qualification rebuttal.

¹⁷¹ We are arguing in favor of the analytical structure, not the merits of the accepted rebuttal evidence, which, in our opinion, was problematic.

¹⁷² *Supra*, footnote 162. While *Etobicoke* was a treatment case (early retirement), the test now applies to both treatment and impact cases with savings clauses. This case is discussed by McIntyre J. in *Bhinder*, *supra*, footnote 4, at pp. 498 (S.C.R.), 144 (C.C.E.L.).

limitation was necessary;¹⁷³ the objective part of the test requires that the limitation be related to the performance of the employment "in that it is reasonably necessary to assure the efficient and economical performance of the job without endangering the employee, his fellow employees and the general public".¹⁷⁴ In *Bhinder*¹⁷⁵ the court found that the employer's "hard hat" rule, although it was *prima facie* discrimination against Sikhs, was justified as a *bona fide* occupational qualification.

When an applicable saving clause is not included in the legislation, *O'Malley* held that the employer must take reasonable steps to accommodate the employee.¹⁷⁶ In this case the Ontario legislation in effect at the time did not attach a saving clause to the prohibition against religious discrimination. The employer dismissed the complainant O'Malley because her religion (Seventh Day Adventist) prevented her from working on Friday nights or Saturdays. The Supreme Court ruled that in the absence of a saving clause, the employer's rule must be "rationally connected to the employment",¹⁷⁷ and also the employer must "take reasonable steps to accommodate the complainant, short of undue hardship: in other words, to take such steps as may be reasonable to accommodate without undue interference in the operation of the employer's business and without undue expense to the employer".¹⁷⁸ While Simpsons-Sears had made some effort to placate through the offer of a part-time job, this was held to be insufficient.

¹⁷³ The subjective element of the test states (*ibid.*, at p. 208):

To be a *bona fide* occupational qualification and requirement a limitation, such as a mandatory retirement at a fixed age, must be imposed honestly, in good faith, and in the sincerely held belief that such limitation is imposed in the interests of the adequate performance of the work involved with all reasonable dispatch, safety, and economy, and not for ulterior or extraneous reasons aimed at objectives which could defeat the purpose of the Code . . .

¹⁷⁴ *Ibid.*

¹⁷⁵ *Supra*, footnote 4, at pp. 499 (D.L.R.), 144-145 (C.C.E.L.)

¹⁷⁶ The duty to accommodate appears to have been introduced in Canadian employment discrimination case law by Professor Cumming in *Singh v. S.I.S.*, *supra*, footnote 162, although R. Juriansz, Survey of Anti-Discrimination Law (1984), 16 Ottawa L. Rev. 117, at p. 129, claims that Cumming, despite an exhaustive survey of American law, actually used an Ontario labour arbitration decision for authority that an employer should make some accommodation. Whatever the genesis of the duty to accommodate, it appears to be now firmly established in Canadian law. While it seems that if a statute contains no saving clause whatsoever, the duty to accommodate will apply to all classifications of discrimination, religion may be a special case. In the U.S. freedom from religious discrimination has been given special statutory protection. It will be interesting to see whether Canadian courts will follow that lead by requiring a higher standard of accommodation in those cases.

¹⁷⁷ *Supra*, footnote 5, at pp. 335 (D.L.R.), 203 (C.C.E.L.); see also "rationally related to the performance of the job" (at pp. 333 (D.L.R.), 200 (C.C.E.L.)).

¹⁷⁸ *Ibid.*, at pp. 335 (D.L.R.), 202 (C.C.E.L.); McIntyre J. continues, at pp. 335 (D.L.R.), 203 (C.C.E.L.):

It is unclear what the requirement of accommodation will mean in practical terms. In *O'Malley* the employer's case at the Tribunal level was argued on the basis of lack of intent, and the issue of hardship was not thoroughly canvassed. The Supreme Court said:¹⁷⁹

There was no evidence adduced regarding the problems which could have arisen as a result of further steps by the respondent, or of what expense would have been incurred in rearranging working periods for her benefit, or of what other problems could have arisen if further steps were taken towards her accommodation. There was therefore no evidence upon which the Board Chairman could have found that such further steps would have caused undue hardship for the respondent and thus have been unreasonable.

This "saving clause/no saving clause" dichotomy did not receive unanimous approval of the Supreme Court. Dickson C.J.C. (Lamer J. concurring), in his lengthy and convincing dissent in *Bhinder*, contends that the duty to accommodate should be imposed even where a saving clause exists. He asserts that a saving clause does not "obliterate the duty to accommodate thereby seriously diminishing the protection from adverse effect discrimination provided in the Act".¹⁸⁰ He also argues that an occupational requirement is not *bona fide* if its application to the employee "is not reasonably necessary in the sense that undue hardship on the part of the employer would result if an exception or substitution for the requirement were allowed in the case of the individual".¹⁸¹

C. Use of Statistical Evidence in Employment Discrimination Cases in Canada

The Supreme Court of Canada's decisions in *Bhinder* and *O'Malley* will undoubtedly encourage the increased use of statistics and statistical evidence in human rights cases in Canada. As recently as 1984, Professor Cumming noted in *Blake v. The Ministry of Correctional Services and Mimico Correctional Institute*,¹⁸² that "[s]tatistical evidence is commonly used in American discrimination cases but has only recently been adduced at Canadian human rights hearings". However, a number of Canadian cases have analyzed statistical data.¹⁸³

Where such reasonable steps, however, do not fully reach the desired end, the complainant, in the absence of some accommodating steps on his own part such as an acceptance in this case of part-time work, must either sacrifice his religious principles or his employment.

¹⁷⁹ *Ibid.*, at pp. 338 (D.L.R.), 206 (C.C.E.L.).

¹⁸⁰ *Supra*, footnote 4, at pp. 487 (D.L.R.), 153 (C.C.E.L.).

¹⁸¹ *Ibid.*, at pp. 488 (D.L.R.), 153 (C.C.E.L.).

¹⁸² (1984), 5 C.H.H.R. D/2417, at p. 2425 (Ontario Board of Inquiry).

¹⁸³ *Hendry v. Liquor Control Board of Ontario* (1980), 1 C.H.H.R. D/160 (Ontario Board of Inquiry); *Officerski v. Peterborough Board of Education* (1980), 1 C.H.H.R. D/33 (Ontario Board of Inquiry); *Ingram v. Natural Footwear* (1980), 1 C.H.H.R. D/59 (Ontario Board of Inquiry); *Dhaliwal and Director, Human Rights Code v. B.C. Timber*

Canadian courts and tribunals are beginning to incorporate the United States courts' experience with the parity and materiality issues. Consider first the parity issue. In *Ingram*,¹⁸⁴ statistics were utilized to show disparities in dismissal rates; in *Dhaliwal*,¹⁸⁵ disparities in hiring rates (cells M and N of Table 1);¹⁸⁶ in *Officerski*,¹⁸⁷ the number of women employed in a particular job was compared to the number of qualified women in the labour market; and in *Blake*,¹⁸⁸ the court compared the percentage of women actually hired with the percentage of women applying for positions (cell M). In *Action Travail*,¹⁸⁹ disparities arose between the percentage of female employees at Canadian National (6.11%) compared to the percentage of females in the Canadian workforce (40.7%) in 1981 (cell A), and between the percentage of women in blue-collar jobs at Canadian National (.7%) compared to the percentage of women in blue collar jobs in the Canadian workforce (13%) in 1981 (cell G). Data were also presented showing low hiring rates of women in non-traditional jobs (cell D).¹⁹⁰ Our review of the Canadian cases suggests that thus far only a limited set of comparisons have been made. We have identified no case, for example, where intra-firm comparisons have been made. As potential plaintiffs realize the full impact of *Bhinder* and *O'Malley* we expect cases will fall in the complete range of cells identified in Table 1.¹⁹¹

While evidence suggests that progress is being made on the parity issue, the materiality issue is more problematic. Of course, concern does not arise in cases involving overwhelming or gross disparities. For example in *Dhaliwal*,¹⁹² the Board examined both hiring rates and employment level without addressing statistical significance:

If one calculates the number of non-East Indian applicants in 1979 through 1981 and compares this with the number of non-East Indians hired, the success rate, by my calculations, appears to have been in the order of 43%, 28% and 14.3% respectively. Taking the comparable figures for East Indians their success rate over the same period was 0%, 6.9% and 11.7%. These figures are for all employees . . . If

Ltd., *supra*, footnote 159; *Marcotte v. Rio Algom*, *supra*, footnote 162; *Carson v. Air Canada*, *supra*, footnote 162; *Action Travail des Femmes and Canadian Human Rights Commission v. Canadian National Railway Company*, *supra*, footnote 7; *Blake v. Ministry of Correctional Services*, *supra*, footnote 182.

¹⁸⁴ *Ibid.*

¹⁸⁵ *Supra*, footnote 159.

¹⁸⁶ *Supra*, p. 670.

¹⁸⁷ *Supra*, footnote 183.

¹⁸⁸ *Supra*, footnote 182. The Board correctly pointed out that this comparison would not have been appropriate where the applications of protected groups were "chilled" by knowledge that there was hiring bias.

¹⁸⁹ *Supra*, footnote 7, at p. 2367.

¹⁹⁰ *Ibid.*, at p. 2368.

¹⁹¹ *Supra*, p. 670.

¹⁹² *Supra*, footnote 159, at p. D/1528.

one eliminates the two East Indian tradesmen hired and looks at entry level jobs alone, the percentage of East Indians hired in 1980 and 1981 drops dramatically.

Again, based on the year end figures the percentage of East Indians in the work force in the plant is very small; 4.91% at the peak, 2.96% in 1980 and 3.06% in 1981, according to my calculations.

Blake is most interesting from a statistical point of view because considerable statistical evidence was presented by expert witnesses for the complainant and because the Board attempted to summarize the parity and materiality issues. The review suggests that Canadian tribunals and courts encounter as much difficulty as United States courts had in the past in distinguishing between the absolute disparity standard and statistical significance. For example, in the excerpt below the Board confuses the concepts of two standard deviations and two percentage points:¹⁹³

The disparity when a neutral employer randomly selects employees from among the applicants is generally less than 2 to 3 standard deviations. A greater disparity gives rise to an inference of discriminatory intent and, if sufficiently large, shifts the burden of proof to the employer to show that discriminatory intent is not the cause of the disparity, or that the statistics inaccurately show the disparity to be greater than it in fact is. . . There is no hard and fast rule that a disparity of greater than 2 percentage points proves discriminatory intent while a disparity of less than 2 may be caused by factors other than sexual discrimination. On the other hand, a disparity of less than 2 may be some evidence of discriminatory intent where there is other evidence of discrimination. The importance of the size of the disparity depends on the cogency and persuasiveness of other evidence in proving or disproving discrimination.

Nonetheless, the Board approved of evidence that differences in male/female interview rates were statistically significant. It also spoke approvingly of the complainant's regression analysis:¹⁹⁴

With his "multiple regression" analysis, Mr. McKie determined that "weight" rather than "sex" was the most significant factor, and that, considering everything, "national origin" and "weight" were the two most significant factors in deciding whether interviews would be given. Once weight is in actuality a factor, there is, of course, a significantly disparate impact upon women in seeking employment, even though there is no formal minimal weight requirement.

Despite the univariate and multivariate statistical evidence, however, the Board found refuge in the gross disparities: "Put another way, 1.1% of female applicants were hired as opposed to 11.58% of male applicants. In my opinion, a *prima facie* case of discrimination because of sex under the *Code* was made out by [these] statistics alone in the instant Inquiry."¹⁹⁵

Action Travail is perhaps representative of Canadian cases in its use of probative comparisons and statistics. While the comparisons made in the case can be assigned to the various cells of Table 1, a considerable

¹⁹³ *Supra*, footnote 182, at p. D/2430.

¹⁹⁴ *Ibid.*, at p. D/2431.

¹⁹⁵ *Ibid.*, at p. D/2432.

effort is required to do so because the Tribunal was not explicit about which particular comparisons were considered probative, why they were selected or which were omitted.¹⁹⁶ Additionally, no attempt was made to examine the materiality (legal significance) of the presented comparisons.¹⁹⁷

In addition to problems relating to the use and interpretation of statistics, there are two further areas of concern. First, the lack of employment data will severely hamper any efforts of tribunals and courts to utilize sophisticated parity comparisons or statistical analyses. This was recognized by the Board in *Hendry*:¹⁹⁸

The LCBO statistical data is incomplete: it does not disclose the proportion of male applicants to female applicants nor their respective success rates in obtaining employment; it does not disclose the annual rate of staff turnover and therefore the opportunities to correct historical imbalances (if the turnover rate is very low, it will take a much longer time to make corrections). Nor did we receive any information on LCBO practices in advertising and recruiting, if any. The minimal statistical information we do have suggests a very large imbalance as between men and women and very little progress in changing that balance. In the absence of any other evidence on the subject, I have concluded that the recruiting practices of the LCBO are largely passive. With little if any active recruitment, it has had little opportunity to affect existing discriminatory patterns.

This statement highlights the understandable frustration of tribunals and courts when the basic statistical data are not available.

Second, the "pre-Act/post-Act" issue is crucial. As noted above, this problem is only beginning to recede in the United States, after more than twenty years. Given the long lead times involved in hierarchical promotion systems, many United States critics claim that several more decades must pass before the issue is finally irrelevant.¹⁹⁹ Obviously it is likely to be much longer before its importance recedes in Canada. The recent rulings in *Bhinder* and *O'Malley* suggest that adverse impact has been in effect in the various jurisdictions since the proclamation of the Acts in these jurisdictions. This implies that employers could be found to have engaged in discriminatory practices if their employment practices had an adverse impact on a protected group since the enactment date. In some jurisdictions this will be taken as many years ago and there will be a relatively minor "pre-Act/post-Act" problem. However, in jurisdictions that have reversed themselves a number of times on the issue of whether or not adverse impact could establish a *prima facie* but rebuttable case, the interpretive vacillations of the courts leave considerable doubt as to the effective date of the introduction of the adverse impact standard.

¹⁹⁶ Of course this responsibility cannot be laid solely at the feet of the Tribunal. The parties to the action have the responsibility for introducing such statistics.

¹⁹⁷ Except to the extent that mention is made of "the inexorable zero"; *supra*, footnote 7, at p. 2369.

¹⁹⁸ *Supra*, footnote 183, at p. D/165.

¹⁹⁹ *Supra*, footnote 61.

Conclusion

Without adverse impact procedures and aggregate statistical evidence, the finding of discrimination depends to a large extent on the subjective interpretation of subjective evidence. Discrimination is difficult to prove in anything but the most overt cases. Therefore, the Supreme Court decisions in *Bhinder* and *O'Malley* are as important in Canada as *Griggs* was in the United States.

As a result of these decisions, Parliament and the provincial legislatures will presumably review their human rights codes and determine whether to include explicitly a requirement of intent and, second, whether to include a saving clause to cover any or all of the protected classes. Undoubtedly, these issues will also be affected by the equality provision of the Charter of Rights and Freedoms.²⁰⁰

A major consequence of *Bhinder* and *O'Malley* will be the increased importance of aggregate statistical evidence in establishing and rebutting *prima facie* cases of employment discrimination. The main purpose of this paper has been to encourage the correct application and interpretation of statistical evidence. We have attempted to perform this task by synthesizing the use of statistics in United States employment discrimination cases over more than twenty-five years. While there are some unresolved issues and statistics have their limitations, an understanding of the United States experience should ensure that Canada does not repeat the mistakes and employ the time, energy and resources the United States has consumed over the years.

The "core" lesson is perhaps almost wholly positive—workable strategies can be developed to measure employment disparity. There are, however, some negative aspects. Measuring and adjudicating issues of employment discrimination is complex and costly—and getting more so. Along with an absolute increase in the number of cases in which statistics will likely become relevant, the level of sophistication required of all parties will undoubtedly rise. Employers, employees, lawyers, judges, and human rights officials will all have to become familiar with a number of methodological issues and problems. It seems clear that the Canadian courts have not yet reached the same level of sophistication and resources as their United States counterparts.

One also must address the quality of existing data. Prior to the recent Supreme Court decisions there appeared to be little incentive for potential plaintiffs to collect data to prove discrimination. Now plaintiffs, in an effort to obtain these data, must confront defendants who may not have collected them or who may have collected them incompletely or inaccurately. In practice, the most important factor is often where the responsibility lies for collecting and making available the relevant personnel data.

²⁰⁰ Part I, Constitution Act, 1981.

It is submitted that there are two possible (non-mutually exclusive) approaches. First, an employer's unwillingness to collect basic statistical data could itself be interpreted by the court as *prima facie* evidence of discrimination. Second, the government could mandate collection of employment data for employers over a specified size. Despite the weaknesses of the Equal Employment Opportunity Commission approach described above, the agency has ensured that employers perform this essential data collection task in the United States.