Credit Scoring and the ECOA: Applying the Effects Test
Credit Scoring and the ECOA: Applying the Effects Test

When Congress enacted the Equal Credit Opportunity Act (ECOA), it announced a federal commitment to combat discrimination on the basis of sex and marital status in credit granting. That commitment was reaffirmed and broadened in 1976 when the ECOA was amended. But difficulties with implementation remain. The goal of the Act is not simply to ameliorate blatant discrimination, but also to facilitate equal access to credit for previously excluded groups by removing practices that have disparate effects. As a result, litigation under the ECOA will raise in the credit context the difficult issues posed by objective screening and the business necessity defense—issues that have been the focus of litigation under Title VII of the Civil Rights Act of 1964. Absent evidence of blatant and deliberate discrimina-

2. Pub. L. No. 93-495, § 502, 88 Stat. 1521 (1974) ("It is the purpose of this Act to require that financial institutions and other firms engaged in the extension of credit make that credit equally available to all credit-worthy customers without regard to sex or marital status.")
   The prohibitions against discrimination on the basis of race, color, religion or national origin are unqualified. In the Committee's view, these characteristics are totally unrelated to creditworthiness and cannot be considered by any creditor. In determining the existence of discrimination on these grounds, . . . courts or agencies are free to look at the effects of a creditor's practices as well as the creditor's motives or conduct in individual transactions. Thus judicial constructions of anti-discrimination legislation in the employment field, in cases such as Griggs v. Duke Power Company . . ., and Albemarle Paper Company v. Moody . . ., are intended to serve as guides in the application of this Act, especially with respect to the allocations of burdens of proof.
Credit Scoring

tion, the ECOA requires judgments about who should bear the cost of eliminating discriminatory effects. This Note argues that placing the burden on creditors, through strict enforcement of the ECOA, is warranted. Strict enforcement will not only help to eradicate discrimination, but will also lead to changes in credit procedures that ultimately advance the creditors’ own interests.

The use of scored objective tests for screening credit applicants makes it unusually difficult to identify and remedy discrimination. These statistically derived, facially neutral screening techniques are often used instead of subjective screening by an individual evaluator; proponents claim that they decrease screening costs and increase predictive accuracy and thus serve to eliminate discrimination. But, as litigation under Title VII has indicated, many problems of discriminatory impact have continued to arise despite the use of objective screening. There remains a basic tension between the central

6. The term “facially neutral” as applied in employment discrimination cases has been used to describe evaluative factors or processes that (1) bear no explicit relationship to group membership, usually defined by an immutable characteristic such as race or sex, but (2) may be sufficiently correlated with such group membership that, absent other explanation, a claim of neutrality in their use is undercut.

7. See Underwood, Law and the Crystal Ball: Predicting Behavior With Statistical Inference and Individualized Judgment, 88 Yale L.J. 1408, 1420-21 (1979) (distinguishing “statistical” and “clinical” predictive techniques).


9. Cf. id. (use of objective criteria avoids unintentional bias often found in subjective employment decisions).

10. See, e.g., Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975) (job placement tests having disparate impact were insufficiently validated); United States v. City of Chicago, 573 F.2d 416, 425-26 (7th Cir. 1978) (reversing finding of no discrimination and remanding for findings of fact regarding whether functions tested in police captain’s promotional exam were part of job function and reflected importance to job). In addition to applying the effects test to objectively scored tests, the courts have applied it routinely in two other kinds of cases involving putative objective criteria used to determine access to jobs. The first kind are cases in which unscored objective criteria of achievement, such as the high school diploma required by employers in Griggs v. Duke Power Co., 401 U.S. 424 (1971), or other factors such as arrest records, convictions, or garnishments, operate as exclusionary devices, see, e.g., Green v. Missouri Pac. R.R., 523 F.2d 1290 (8th Cir. 1975) (convictions); Gregory v. Litton Sys., Inc., 316 F. Supp. 401 (C.D. Cal. 1970), modified, 472 F.2d 631 (9th Cir. 1972) (arrests); Johnson v. Pike, 332 F. Supp. 490 (C.D. Cal. 1971) (garnishments). The second kind of case involves the use of unstructured criteria to allocate opportunities. See, e.g., Hazelwood School Dist. v. United States, 433 U.S. 299, 302 (1977) (only general guidance given to principals to hire most competent teachers; intangibles such as “personality, disposition, appearance, poise, voice, articulation, and ability to deal with people” counted heavily). The courts have also addressed the issue of whether the effects test applies to the use of sex-related criteria in the allocation of job benefits. See Nashville Gas Co. v. Satty, 434 U.S. 136 (1977) (denial of pregnancy disability benefits and back pay to workers on pregnancy leave not discrimination but loss of seniority for same time off is sex discrimination); General Elec. Co. v. Gilbert, 429 U.S. 125 (1976) (denial of pregnancy health and disability benefits not sex discrimination at least when men do not receive health benefits of greater value than women when pregnancy excluded from coverage).
principle of equality, which demands that scarce and valuable resources be allocated according to individual qualification, and the idea that statistical prediction, derived from broad group characteristics, is a valid means for making such a determination.

This Note will argue that judicial holdings with respect to statistical prediction in Title VII cases articulate principles that can be extended to consumer credit-screening transactions subject to the ECOA. In Title VII cases, which involve discrimination in employment, the judiciary has not deferred to the appearance of objectivity and has not relied wholly on rigid, technical methods of statistical analysis. Instead the courts have judged the screening transaction as a whole. By examining the standards implicit in Title VII cases, this Note derives an analogous test for credit scoring and determines the stringency with which that test should apply. Part I describes consumer credit scoring and analyzes the congressional purpose that underlies the "effects test" that was incorporated in the ECOA Amendments. Part II describes the effects test as it is used in the employment context and identifies the similarities and differences between employment and consumer credit. Finally, Part III proposes a method for applying the test to the credit-scoring context.

I. Consumer Credit Decisionmaking and the Equal Credit Opportunity Act

The process of consumer credit screening has three interrelated phases—recruitment or solicitation, evaluation, and monitoring and feedback. Although discrimination is found in all three phases of the process, the terms of the ECOA focus primarily on evaluation. The success of the ECOA depends on an understanding of credit evaluation and the sources of discrimination in that process.

11. See pp. 1462-69 infra.
12. Recruitment or solicitation refers to all the means by which a creditor attracts applications. The creditor's general advertising strategy is one factor that influences the applications it receives. See R. Cole, CONSUMER AND COMMERCIAL CREDIT MANAGEMENT 351 (5th ed. 1976).
13. Evaluation is the process by which creditors decide to accept or reject an application. See generally id. at 245-847 (discussing evaluative factors in credit investigation, means and shortcomings of investigation and information verification, and techniques of evaluation).
14. Monitoring and feedback lead to two types of actions by the creditor. First, the creditor may terminate or pursue collection strategies for accounts that fail to maintain a good payment record. Id. at 363-87. Second, the creditor uses past experience with accounts to guide his subsequent solicitation and evaluation procedures. Id. at 308-22.
15. The Act prohibits discrimination with respect to any aspect of a credit transaction, see 15 U.S.C. § 1691(a) (1976); however, most of its specific provisions and most provisions of the implementing regulations issued by the Federal Reserve Board, see Regulation B,
Credit Scoring

A. Consumer Credit Evaluation

Credit evaluation is a two-tiered process. First, the creditor screens out those applicants who, from the point of view of the creditor, are bad risks. They include people who are apparently unwilling or unable to meet their obligations and also those against whom the creditor would be unable to assert an effective claim in the event of default.\(^6\) The screening process may require verification of information provided by the applicant or further investigation through other sources.\(^7\)

During the second stage of the process, the creditor evaluates his applicant pool to achieve a portfolio of debtors with an expected likelihood of repayment consistent with that creditor's desired risk level.\(^8\)

Creditors employ two basic methods to screen credit applicants: they may rely on subjective evaluations of individual applicants\(^9\) or on standardized screening devices that use statistical analysis of the creditor's past experience. The statistical screening devices, which are generally referred to as credit scoring,\(^2\) are most often used by

12 C.F.R. § 202 (1978), pertain to the information that a creditor may request of an applicant, the use to which permissible information may be put, and a creditor's duties to a rejected applicant.

16. See R. Cole, supra note 12, at 245-307. In addition to typical screening of applicants for positive indicators of financial ability and a satisfactory history of meeting obligations, the creditor attempts to screen out judgment-proof or fraudulent applicants who, it is reasonable to infer, cause the most substantial losses in the event of default. Cf. Ross, The Credit Card's Painful Coming-of-Age, Fortune, Oct. 1971, at 111 (unsolicited mailing of bank cards led to substantial losses from fraudulent use).

17. See R. Cole, supra note 12, at 267-307 (discussing direct investigation with references provided by applicant and indirect verification through credit-reporting agency).

18. Id. at 308-22 (discussing means of risk evaluation including credit officer appraisal, credit grading, and credit scoring). Creditors, finding it uneconomical to vary prices according to differences in contractual terms and perceived risks, tend to establish a “house” interest rate and lend only to those who meet a minimum risk level; consequently interest rates vary more among lenders than they do among the customers of one lender. National Commission on Consumer Finance, Consumer Credit in the United States 112-13 (1972). The overall effect is stratification of the credit industry so that borrowers perceived as high risks obtain credit from creditors, such as finance companies, who offer credit at higher interest rates. See R. Cole, supra note 12, at 102-03.


20. Id. at 319-22. Credit scoring, which is statistically derived, is thus distinguished from credit grading, which merely systematizes the subjective standards chosen by the creditor. Id. at 312-22; see 12 C.F.R. § 202.2(p)(1) (1978):

Empirically derived credit system.

(i) The term means a credit scoring system that evaluates an applicant's credit-worthiness primarily by allocating points . . . to key attributes describing the applicant and other aspects of the transaction. In such a system, the points . . . assigned to each attribute, and hence the entire score:

(ii) Are derived from an empirical comparison of sample groups or the population of creditworthy and noncreditworthy applicants of a creditor who applied for credit within a reasonable preceding period of time; and

(iii) Determine alone or in conjunction with an evaluation of additional information about the applicant, whether an applicant is deemed creditworthy.

1453
creditors handling a large volume of consumer debt and screening numerous applicants. Therefore, a discriminatory credit-scoring system will have a more significant impact than a single subjective process. In addition, credit scoring creates more difficult analytical problems when the law seeks to end discrimination in the credit industry.

In credit scoring, the applicant is scored on a combination of characteristics to determine whether he meets the creditor’s predetermined risk level. A scoring scheme typically employs six to twelve variables relating risk of nonpayment to attributes possessed by the applicant or associated with the transaction. These variables may involve the terms of the loan, credit and financial references, income, job status, residence, or a variety of other factors. The score that an applicant receives on each variable is derived from the probability of bad payment associated with that variable in the creditor’s pool of past credit recipients.

21. Although the goals of uniform treatment of applicants and systematic prediction of risk are no doubt factors in a creditor’s decision to adopt a credit-scoring system, efficiency, scale, and computer capabilities may play a larger role in that decision. See, e.g., R. Biborosch, Numerical Credit Scoring (pt. 1) (speech at Installment Credit Conference of American Bankers Association), reprinted in CREDIT WORLD, June 1965, at 6-9 (First Pennsylvania Banking and Trust Co. first moved to credit-grading procedure and systematic numerical scoring to accommodate substantial volume of postwar business at time when trained and experienced credit personnel were unavailable or drew salaries higher than bank budgeted); Wells, New Customer Credit Pointing System (Jan. 10, 1963) (presentation at Second Consumer Credit Symposium, University of Pennsylvania), reprinted in Numerical Pointing Plans for Evaluating Consumer Credit Risks (pamphlet on file with Yale Law Journal) (describing principles of credit-scoring system developed by mail-order business to compensate for lack of face-to-face screening capabilities and to avoid time and costs entailed in ordering credit reports; generally recommending its use for volume operations in retail credit field). See generally STANFORD RESEARCH INSTITUTE, THE ROLE OF RISK CLASSIFICATION IN PROPERTY AND CASUALTY INSURANCE 96 (Supp. 1976) (credit scoring adopted as result of economic forces calling for lowest possible costs and highest possible volume and political pressure aimed at avoiding discrimination).


23. Although the numerical credit-scoring scheme may be developed with any of several mathematical techniques, the fundamental procedure for deriving such a scheme is a three-step process. The first step is to determine on the basis of a sample of past credit users which variables or factors retained in the file on credit users demonstrate a mathematically meaningful relationship with the probability of failure to make timely voluntary payment. The second step is to devise a scoring scheme employing the variables that proved to be predictive. A weight for each variable and a score for each possible category subsumed by the variable are assigned based on the predictive values obtained. Combining the values of each attribute and any weighting given to the variables, the numerical scoring scheme assigns a single, typically additive score for any combination of attributes. This score represents the probability that a credit user in the creditor’s file who possesses the combined attributes is a good credit risk. Ideally that probability is tested
Credit Scoring

A credit-scoring scheme reveals explicitly the factors considered in decisionmaking. This would seem to make discrimination on the basis of the applicant's sex or race more difficult than in subjective screening by a credit evaluator.\textsuperscript{24} On the other hand, the complexity of credit-scoring systems increases the chances that they will intentionally or inadvertently be used to impose disadvantages on creditworthy members of statutorily protected classes,\textsuperscript{25} notably racial minorities and women. In many credit-scoring schemes, the creditor's preference for certain kinds of credit users,\textsuperscript{26} or his intuition about against a random split sample of previous credit users or against a sample of subsequent credit users.

The third step in credit scoring is estimating the point below which the probable losses associated with bad accounts are equal to or exceed the aggregate desired gain from good accounts. A scoring scheme may be used as a simple elimination device or it may be constructed so that the lowest-scoring applicants are automatically rejected, the highest scoring are automatically accepted, and the middle range are held for closer examination. See Hsia, \textit{Credit Scoring and the Equal Credit Opportunity Act}, 30 Hastings L.J. 371, 375-77, 387-404 (1978); Myers & Forgy, \textit{supra} note 22; Smith, \textit{supra} note 22.

\textsuperscript{24} It has been argued that objective, empirically derived systems are superior to judgmental systems, and preferable from the viewpoint of discrimination law, because a credit officer's judgment is based on imperfect analysis and distorted recollection of information that is, in addition, likely to be out-of-date. Hsia, \textit{supra} note 23, at 372-75. However, a statistical device like credit scoring can lead to systematic mismeasurement, see pp. 1455-58 infra, and is not necessarily preferable on policy grounds. For discussion of policy considerations associated with the decision to use statistical or clinical prediction, see Underwood, \textit{supra} note 7, at 1420-32.

\textsuperscript{25} See \textit{Hearings on H.R. 14856 and H.R. 14908 Before the Subcomm. on Consumer Affairs of the House Comm. on Banking and Currency}, 93d Cong., 2d Sess. 31-32 (1974) (statement of Jeffrey M. Bucher, member Federal Reserve Board) (seemingly innocent inquiry regarding address may be used by creditor to discount or avoid neighborhoods with high minority population) [hereinafter cited as 1974 \textit{Hearings}]; Chandler & Ewert, \textit{Discrimination on the Basis of Sex under the Equal Credit Opportunity Act} (1976), discussed in \textit{Proceedings of ABA National Institute on Consumer Credit}, 33 Bus. L. 1086, 1090-1104 (1978) (construct of credit-scoring systems with four alternative models—one aggregating male and female performance, one distinguishing sex with dichotomous variables, one based on males only, and one based on females only—reveals that correlation of female performance with predictive variables is different from that of males and that aggregate model underpredicts good performance of females) (on file with \textit{Yale Law Journal}); Eisenbeis, \textit{Problems in Applying Discriminant Analysis in Credit Scoring Models} (1977) (Board of Governors of the Federal Reserve System Staff Economics Study No. 94) (shortcomings of credit-scoring models include serious problems of bias resulting from estimation based on truncated sample).

\textsuperscript{26} Such preferences may be expressed in subtle ways that have discriminatory effects. Some creditors who use occupation as a variable in credit-scoring schemes score occupation by subgroupings such as professional, semiprofessional, managerial, clerical, laborer, etc. Although some occupations indisputably fall into one subgrouping or another, many occupations could be placed in more than one. For example, an accountant or a nurse might be defined as professional or semiprofessional, and an administrative assistant could be managerial or clerical. Telephone interview with Sally Gold, attorney, Division of Credit Practices, Bureau of Consumer Protection, Federal Trade Commission, Washington, D.C. (Mar. 30, 1979). The decision to place an occupation in professional or managerial categories rather than semiprofessional or clerical, is in effect a decision to benefit the particular occupation by attributing to it the aggregate score that the category sub-
the shortcomings of his data, limits the empirically derived measures. Often the creditor does not define the variables in the system to account for the full range of possibilities.

Even if the creditor faithfully relies on his available data, the derived scores may not be correct predictors of creditworthiness for members of protected classes. Often, particular predictive variables correlate differently with repayment behavior for white males than for other classes of applicants. This effect will be fairly obvious when a variable that is valued highly is possessed in low proportion by protected classes or vice versa. But the differential may not always be easy to perceive. In either case, inadequate prediction for protected classes is the result because patterns of geographic dis-

suming the higher status occupations is likely to achieve. For a description of techniques by which constraints can be imposed on empirical derivation, and variables can be re-grouped and reordered in computer data analysis, see generally N. Nie, C. Hull, J. Jenkins, K. Steinbrener & D. Bent, Statistical Package for the Social Sciences 89-126 (5th ed. 1975).

In addition to expressing preferences at the stage of subcategorizing occupations, a creditor who exercises discretion when interpreting individual applications might decide, for example, whether an "administrative assistant" fits managerial or clerical subcategories by considering the sex of the applicant which might be discerned from the applicant's name.

27. The foundation of such an intuition may be sound or unsound. Two reasons for employing constraints arise frequently. First, there may be so few previous cases in a subcategory that it is statistically unreliable. See generally H. Blalock, Social Statistics 162 (2d ed. 1972) (with very small samples, extreme results can occur quite frequently by chance). Second, the creditor's sample may not include a portion of the population that is now applying. Hsia, supra note 23, at 389-91, 391 n.75, 393-98 (problem of estimation and technique for compensation with sample truncated by exclusion of previously rejected applicants; requirement that accounts achieve given age to yield useful information; problem of sample rendered unrepresentative by passage of time and change in demographic factors); Eisenbeis, supra note 25, at 20-22 (problem of estimation with truncated sample).

28. The most prominent example is the use of zip codes or residence regions that correspond to racial residential patterning and the consequent assignment of a low score to minority areas. See, e.g., Robinson, Zip Code Factor in Penney's Credit Checks, Raleigh Times, June 28, 1977, at 1-A; Wilcox, U.S. Investigates Credit Card Redlining, Detroit News, Aug. 17, 1977, at 3-A, col. 1.

29. See Chandler & Ewert, supra note 25, at 6-10 (scoring model shows different predictive weights for model scoring system based on female subpopulation only compared with model based on male and female populations combined and compared with model derived from males only).

30. In this context, "inadequate prediction" refers to a scoring system's inability to identify the same percentage of successes in each group of the population. It is not overall accuracy of a scoring system compared with other procedures that is at issue in this discussion. The issue here is the legitimacy of using systems that may be expected to systematically predict a lower proportion of successes for previously disadvantaged groups than for historically advantaged groups. However, if a system is highly inaccurate, its legitimacy may be further questioned on that ground alone. See generally Underwood, supra note 7, at 1410-12.
Credit Scoring

tribution, and financial practices among protected groups differ from the patterns of those characteristics among white males. For example, although home ownership may be a good indicator that a given applicant is a good risk, holding financial ability constant, more white men will have acquired ownership of personal residences than will women or minorities. As a result, home ownership probably identifies a larger proportion of financially responsible white males than of financially responsible women and minorities.

Inadequate prediction of female or minority credit performance may be amplified in two ways by a creditor who has a history of excluding women or minorities. First, if the creditor relies on a scoring scheme based on a sample composed predominantly of white, male credit recipients, he penalizes minority and female applicants for not possessing the positive predictors for white males that are


32. Goulet, Credit Potential of Women, 9 J. Consumer Credit Management 102, 102-06 (1978) (women may be expected to have different credit demand schedules and to differ in relevant attributes; for example, because many women time credit-market activity to coincide with family responsibilities, their work careers average 10 years less than men and their entry into market tends to occur at later age).

33. Although there appears to be no systematic data on patterns of financial practice and credit use according to sex and race, evidence on sex and race discrimination in the credit industry indicates that credit-history characteristics will vary considerably according to sex and race. Compared to white males, women and minorities will have borrowed less, from sources of lesser quality, and usually at higher interest rates, regardless of financial capability or other indicators of financial responsibility. See, e.g., Comptroller of Currency, Fair Housing Lending Practices Pilot Project (1975), reprinted in Hearings on S.381, S.1900, S.1927, S.1961 and H.R. 5616 Before the Subcomm. on Consumer Affairs of the Senate Comm. on Banking, Housing and Urban Affairs, 94th Cong., 1st Sess. 481-526 (1975) [hereinafter cited as 1975 Senate Hearings] (disparities in rejection rate indicated for minorities, as compared to nonminorities, by tables with variables of race, income, and assets); National Commission on Consumer Finance, supra note 18, at 151-60 (clear anecdotal evidence on sex discrimination; equivocal evidence on race); Gates, Credit Discrimination Against Women: Causes and Solutions, 27 Vand. L. Rev. 412 (1974) (survey of credit discrimination encountered by women and evaluation of factors that must be confronted in finding solution to discrimination).

34. See Chandler & Ewert, supra note 25, at 8-9. Using a sample drawn from 2,000 credit applicants to a large metropolitan bank issuing credit cards, Chandler and Ewert found that 15.7% of the females in the sample reported owning their own homes, as compared with 24.9% of the males. Id. Of those who possessed their own homes, 89% of the females showed good credit performance and only 74% of the males had good performance. Id. These results would indicate that women who own their own homes are a very good risk. It indicates nothing about the performance of women who do not own their own homes or who do not possess other attributes that are positive for white males. For example, Chandler and Ewert found that although more women in the sample (58.8%) were classed as low income than men (23.6%), the percentage of low-income females who had good credit performance (60%) was higher than the percentage of low-income males who had good credit performance (49%). Id.
difficult and time consuming to acquire. Second, if the creditor defines, for estimation purposes, the dependent variable "uncreditworthy" to include not only those individuals in the sample who have met some criteria of bad payment behavior but also those who have been rejected in past applications for credit and, therefore, never had an opportunity to perform, this creditor, in effect, assigns negative value to attributes possessed in high proportion by minorities or women regardless of the actual correlation between these attributes and payment behavior.

On the basis of the foregoing factors, we may infer that objective credit-scoring devices presently used by many creditors will not identify creditworthy members of disadvantaged groups as effectively as they identify creditworthy white males. The particular attributes associated with group membership, which necessarily affect measurement devices, make this outcome virtually certain. As a result, enforcement of the ECOA raises problems, like those faced in other contexts, that concern structuring scoring systems properly to treat all equally qualified applicants equally.

35. The fact that scoring is necessarily derived from samples of past credit users has always been a concern in credit prediction. See, e.g., D. Durand, supra note 22, at 7; Hsia, supra note 23, at 389-91 (sampling); id. at 393-98 (techniques to adjust for lack of information on rejects); Smith, supra note 22, at 333, 357-38. Hsia notes that creditors have a variety of possible responses to this problem. One is to assume that all rejected applicants were noncreditworthy and treat them for estimation purposes as if they were bad accounts. Hsia, supra note 23, at 394. Obviously this approach amplifies the negative value of those traits that are characteristic of rejected applicants, some of whom would not have shown bad performance. A second approach described by Hsia is to augment derived scores by the probability that previously accepted applicants would receive that score. Id. at 397-98. This would appear to perpetuate previous discriminatory rejection if it assigned even greater positive value to attributes possessed in low proportion by previously excluded women and minorities.

36. Defining "equal qualifications" is a central issue in discrimination law. This issue has received sophisticated treatment in professional literature on scored predictive testing in employment and education. The primary problem arises when the qualifying standards are more accurately predictive of the ultimate performance of members of one demographic group than of the performance of members of another demographic group. When tests do not predict as accurately for one group as for another, an individual member likely to succeed is less likely to be scored as a potential success than is a member of the group for which the test predicts accurately. Fairness to the individual implies that it is discriminatory to him, because of his group membership, to adhere strictly to the test results in judging his likelihood of success. But adjusting test scores to admit more potentially successful members of the group for which the test does not predict as accurately also entails admitting more potential failures, who will, when opportunities are limited, take an opportunity that could have been given to an individual for whom the test predicted success more accurately. See, e.g., Cole, Bias in Selection, 10 J. Educ. Measurement 237 (1973) (discussing various proposed models of fairness and proposing conditional probability model based on equal opportunity for potentially successful applicants); Thorndike, Concepts of Culture-Fairness, 8 J. Educ. Measurement 63 (1971) (alternative definition of fair test is test set to admit proportion of two groups reaching specified level of criterion performance). The desirability of a given quantitative solution
Credit Scoring

B. Equal Credit Opportunity Act Amendments

The purpose of the ECOA was to eliminate discrimination in the granting of consumer credit and to afford equal access to credit for members of excluded groups. The drafters of the Act made a conscious decision to use the term “discriminate” without defining or qualifying it. They were concerned with more than just intentional discrimination and therefore chose to allow courts to assess the subtleties of credit practices and their impact.

The effects test was incorporated in the 1976 Amendments to the ECOA in order to make it clear that the Act covered not only discriminatory motivation but also the discriminatory effects of facially neutral actions. This action indicates a congressional intent that for fair screening depends on the manner in which the two groups differ in relation to the test. See Einhorn & Bass, Methodological Considerations Relevant to Discrimination in Employment, 75 PSYCHOLOGICAL BULL. 261 (1971) (differences between subgroups with respect to test-criterion correlations, criterion means and variances, and differences in standard errors of estimates are all relevant in avoiding unfair discrimination). Not surprisingly, it has been argued that fairness in test selection cannot be reduced to completely mechanical procedures. Darlington, Another Look at “Cultural Fairness,” 8 J. EDUC. MEASUREMENT 71 (1971) (if conflict arises between two goals of maximizing test’s validity and minimizing test’s discrimination against certain cultural groups, then subjective, policy-level decision must be made concerning relative importance of two goals).

Discrimination law, using language eschewing quotas, has shown reluctance to embrace strategies based on quantitative standards to avoid disparate outcomes. Any quantitative approach, because it is necessarily based on group membership, does not respond to the concern for individual rights that is the foundation of the law. The value placed on individual autonomy, and the societal interest in rewarding individual achievement, ground the legal preference for decisionmaking that is responsive to individual action. Cf. Fiss, A Theory of Fair Employment Laws, 38 U. CHI. L. REV. 235, 240-41 (1971) (racial criterion unfair because race is not accurate predictor of productivity and is beyond individual control); Underwood, supra note 7, at 1454-42 (policies favoring individual autonomy as basis for evaluating use of statistical prediction).

See note 4 supra; 1975 Senate Hearings, supra note 33, at 47-48 (statement of Werner H. Kamarsky, commissioner, New York State Division of Human Rights) (emphasizing that discriminatory consequences, not just motivation, would be covered by Act under effects test).
the courts employ the same factfinding strategies in credit cases as they do in Title VII employment cases. Congress clearly empowered the courts\(^4\) to decide in each case whether a particular method of credit decisionmaking affords an equal opportunity to any creditworthy individual.\(^4\)

The courts will encounter complicated legal and factual issues when they apply the effects test to cases alleging discrimination based on objective credit-scoring systems. When such objective measures are used, it is difficult to distinguish between genuinely valid prediction and inherent bias derived from the nature of the variables measured and the scoring instrument. This problem was raised during the congressional hearings concerning the 1976 Amendments. Industry spokesmen contrasted credit scoring with subjective screening and argued that because a statistical device does not embody subjective intent, it is necessarily unbiased.\(^4\) They also produced evidence that credit scoring aids prediction of aggregate credit performance and improves the overall profitability of the creditor's portfolio.\(^4\)

Some congressional subcommittee members and consumer advocates, on the other hand, were seriously concerned that aggregate measurement techniques "were inherently discriminatory in that

---

43. In addition, the Act empowers the Board of Governors of the Federal Reserve Board to issue regulations that achieve this same purpose. 15 U.S.C. § 1691(b) (1976). When it promulgated Regulation B, 12 C.F.R. § 202 (1978), the Board proscribed certain facially neutral evaluative factors that are expected to have a disparate effect on protected classes. See id. § 202.6(b)(4) (proscribing evaluation by creditor on basis of whether telephone in applicant's own name); id. § 202.6(b)(5) (proscribing creditor discounting of part-time income). However, the Board pointed out early in the hearings on the ECOA that it could not be assumed that use of certain factors—for example, geographic regions—necessarily discriminated against a protected class, and asserted that the discriminatory nature of many factors could only be determined by the courts on a case-by-case basis. See 1974 Hearings, supra note 25, at 32 (statement of Jeffrey M. Bucher, member, Federal Reserve Board).

44. Throughout the hearings leading to the Act and the 1976 Amendments, congressional proponents stressed that the purpose of the ECOA was to provide equal access to creditworthy persons based on individual willingness and ability to repay. At the same time it was acknowledged that the purpose of the Act was not to require creditors to extend credit to uncreditworthy individuals or to forego considering ability to collect in the event of default. See, e.g., 1974 Hearings, supra note 25, at 15.

45. Testimony asserting the desirability of credit scoring invoked its objective quality and likened it to actuarial prediction in casualty insurance. See id. at 61 *passim*; 1975 Senate Hearings, supra note 33, at 227 *passim*; Hearings Before Subcomm. on Consumer Affairs of the House Comm. on Banking, Currency and Housing, 94th Cong., 1st Sess. 80 *passim* (1975) [hereinafter cited as 1975 House Hearings]; but cf. id. at 37 (statement of Jeffrey M. Bucher, member, Federal Reserve Board) (contrasts between credit prediction and death and casualty prediction indicate that credit scoring is less sophisticated and less statistically reliable).

46. See, e.g., 1974 House Hearings, supra note 25, at 438-40; 1975 House Hearings, supra note 45, at 87-95; 1975 Senate Hearings, supra note 33, at 440-45; Senate Report, supra note 4, at 6.
they saddl[e] each applicant with the statistical characteristics of similar prior applicants. The effects test responds to this concern, but, without further specification, does not resolve the problem in the credit-granting context: it remains for the courts to give concrete meaning to the ECOA standards for credit scoring.

II. The Effects Test and Credit Scoring

The existing case law regarding the effects test is restricted almost entirely to the area of employment discrimination prohibited by Title VII. It is natural to look to this case law for guidance in applying the effects test to credit scoring. But these cases are relevant only as analogies. Application of the test must be guided by the nature of the underlying transaction and by the social goals pertaining to that transaction. The factors that are salient in applying the effects test are not the same for credit and employment decisions.

47. Senate Report, supra note 4, at 6; see 1973 Senate Hearings, supra note 33, at 455 (Sen. Biden) (expressing concern that people are “falling into a general category which is used in the point scoring analysis when they really are an exception to that category, and there is no way they can get out from under it”).

48. Owing to the newness of the Act and the time required to bring a disparate impact suit involving complicated statistical proof to trial, no ECOA suit challenging a credit-scoring scheme on effects-test grounds has been fully tried. At least one case has raised an effects-test issue. Carroll v. Exxon Co., 434 F. Supp. 557, 563 (E.D. La. 1977) (cannot conclude as matter of law that inquiry about number of dependents has effect of discriminating against unmarried persons). In addition, the Federal Trade Commission has pressed two cases involving discriminatory-effects issues, along with other issues, to settlement. United States v. Federated Dep’t Stores, Inc., No. 78-750 (S.D. Ohio, settlement filed Nov. 16, 1978); In re Aldens, Inc., 5 CONS. CRED. GUIDE (CCH) ¶ 98,065 (FTC Feb. 14, 1978) (consent order).

49. See, e.g., Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975); Griggs v. Duke Power Co., 401 U.S. 424 (1971). The lower courts have applied a similar concept in public employment cases litigated under the equal protection clause and 42 U.S.C. §§ 1981, 1983 (1976). See, e.g., Castro v. Beccher, 459 F.2d 725 (1st Cir. 1972) (police hiring); Arnold v. Ballard, 390 F. Supp. 723 (N.D. Ohio 1975) (same). But see Harper v. Mayor & City Council of Baltimore, 359 F. Supp. 1187, 1204 n.41 (D. Md. 1973) (fire department employment) (cases under 42 U.S.C. § 1981 involve less stringent prohibition of nonintentional discrimination than cases under Title VII because employer, upon showing business necessity, is not further required to adopt less discriminatory alternative). However, this interpretation was modified when the Supreme Court held in Washington v. Davis, 426 U.S. 229 (1976), that the plaintiff must demonstrate discriminatory intent under the equal protection clause. The holding’s practical effect on public-employment cases was mitigated because Title VII was extended to cover public employers in 1972. See Equal Employment Opportunity Act of 1972, Pub. L. No. 92-261, 86 Stat. 103 (1972) (codified at 42 U.S.C. §§ 2000e to 2000e-14 (1976)). Courts have in some instances applied an effects-test concept to statutorily grounded housing discrimination cases. See generally Comment, Applying the Title VII Prima Facie Case to Title VIII Litigation, 11 HARV. C.R.-C.L. L. REV. 128 (1976). Thus, the courts may be adopting a distinction between statutory discrimination laws and the constitutional requirements of the equal protection clause.
Nonetheless, the development of the effects test in employment law provides a useful starting point for analysis of the test’s application to credit scoring. Close attention must be paid to the ways in which credit scoring and employment screening resemble one another, and to the ways in which they differ; only in this fashion will it be possible to develop an application of the effects test to the credit-scoring process that is responsive to the terms of the Act and the goals of discrimination law.

A. The Effects Test in Employment Cases

The effects test under Title VII operates to shift the burdens of producing evidence and of persuading the court. The plaintiff has the initial burden of proving the prima facie case that a facially neutral procedure or standard has a disparate impact on a protected class. If the adverse impact is shown, the defendant then has the burden of persuading the court that this procedure or standard is based on business necessity.51 Even if the defendant shows a legitimate business purpose for the practices in question, the availability of a less discriminatory means to achieve the same legitimate goals may support a finding of discrimination.52


51. Such presumptions are created to reflect judicial evaluation of probabilities and a party’s superior access to the proof, and to effectuate policies by creating a preferred litigant. See McCormick’s Handbook of the Law of Evidence §§ 337, 343 (E. Cleary 2d ed. 1972); James, Burdens of Proof, 47 VA. L. REV. 51, 61 (1961); cf. Underwood, Thumb on the Scales of Justice, 86 YALE L.J. 1299 (1977) (burdens of persuasion in criminal cases).

52. See Albemarle Paper Co. v. Moody, 422 U.S. 405, 425 (1975) (dictum) (“It remains open to the complaining party to show that other tests or selection devices, without a similarly undesirable racial effect, would also serve the employer’s legitimate interest in ‘efficient and trustworthy workmanship’”); United States v. Bethlehem Steel Corp., 446 F.2d 652, 663 (2d Cir. 1971) (when practices perpetuate past discrimination, necessity connotes irresistible demand; if safety and efficiency can be served by reasonably available alternative system with less discriminatory effects, then present policies yielding disparate effects cannot continue); Robinson v. Lorillard Corp., 444 F.2d 791, 798 & n.7 (4th Cir.), cert. dismissed, 404 U.S. 1006 (1971) (test is whether overriding business purpose for practice necessary to safe and efficient operation and no acceptable practices with lesser racial impact exist).

53. See Green v. Missouri Pac. R.R., 523 F.2d 1290, 1298 (8th Cir. 1975) (alternative holding) (refusal to hire anyone with criminal record discriminatory because insufficient business necessity; individual consideration of significance of conviction is less restrictive means); Pettway v. American Cast Iron Pipe Co., 494 F.2d 211, 245 (5th Cir. 1974) (alternative ground) (availability of validated reading test as less restrictive means); United States v. St. Louis-S.F. Ry., 464 F.2d 301, 309 (8th Cir. 1972), cert. denied, 409 U.S. 1116 (1973) (individual testing of abilities is nondiscriminatory alternative to seniority system for
Credit Scoring

1. The Prima Facie Case

In the early effects-test cases in employment, the facially neutral decision procedures were fairly obvious substitutes for previous, explicitly discriminatory practices. As the practices that were evaluated became more subtle, the courts refined the standards of quantitative and qualitative proof of discrimination. The quantitative method involves comparing the percentage of protected group members among those hired after the effective date of the Act with the percentage belonging to the group in the relevant population. The relevant population can be defined as either the applicant pool or ensuring safety; United States v. Bethlehem Steel Corp., 446 F.2d 652, 662-64 (2d Cir. 1971) (ground for remedy) (government proposed remedy as less restrictive means for promotion system); Robinson v. Lorillard Corp., 444 F.2d 791, 799-800 (4th Cir.), cert. dismissed, 404 U.S. 1006 (1971) (alternative ground) (availability of less discriminatory alternative promotion system one factor in rejection of business necessity argument); B. SCHLEI & P. GROSSMAN, EMPLOYMENT DISCRIMINATION LAW 134-35 & n.15 (1976). See generally Note, Business Necessity under Title VII of the Civil Rights Act of 1964: A No-Alternative Approach, 84 YALE L.J. 98 (1974) (rejecting "business purpose," "no perpetuation," and "balancing" tests for judicial response under Title VII to employment practices having disparate impact; adopting "no alternative" theory that focuses on two questions: (1) Does practice serve legitimate business purpose? (2) Is alternative practice available that will promote business purpose equally well with lesser disparate impact on blacks?).

One commentary has speculated that the "no alternative" showing and the question of which party has the burden of proving the availability of alternatives will be crucial in applying the effects test to the credit context. See Geltzer & Feldman, An Initial Analysis of the Impact of Revised Regulation B on the Retailer and Bank Creditor, 33 BUS. LAW. 47, 49 (1977). Although Albemarle Paper Co. v. Moody, 422 U.S. 405, 425 (1975), in dictum seems to suggest that the plaintiff bears the burden of persuading the court that a less discriminatory alternative exists, some courts state that the defendant has the burden of proving that no less discriminatory practices are available. Wallace v. Debron Corp., 494 F.2d 674, 677 (8th Cir. 1974) (employer must show no acceptable alternative to garnishment policy); Officers for Justice v. Civil Serv. Comm’n, 395 F. Supp. 378, 383-84 (N.D. Cal. 1975) (alternative ground) (defendant did not show that there was no less discriminatory alternative to physical agility test); see 29 C.F.R. § 1607.3 (1978) (EEOC guidelines placing burden on employer to demonstrate that no suitable alternative with lesser impact available).


55. When the disparity is extreme, instead of limiting the comparison to post-Act employment, the courts have compared the composition of the employer’s entire work force with the composition of the relevant population. See, e.g., International Bhd. of Teamsters v. United States, 431 U.S. 324, 337-38 & n.17 (1977) (comparison of employer’s work force with general population appropriate when disparity great and employer’s post-Act hiring practices represent no significant change).

56. Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975) (prima facie case primarily shown by comparison of percentage of actual black applicants who passed test and actual white applicants who passed test).
the general population in the region,\(^5\) or both.\(^5\) If the comparison shows a statistically significant difference, a prima facie case has been established.\(^5\) To ameliorate the stringency of this test as it was initially stated, the courts have recognized two refinements in the use of statistical evidence. First, the universe for statistical comparison, when based on general-population data, has been expressly limited to qualified persons in the relevant geographic region.\(^6\) Second, an employer may defend against a prima facie showing by producing evidence that the basis of comparison is skewed against him for

\(^5\) See Hazelwood School Dist. v. United States, 433 U.S. 299, 308 & n.13 (1977) (general population of qualified teachers in St. Louis County is proper comparison for post-Act hiring rate but defendant entitled to rebut statistics by showing that such population comparison artificially inflated); Griggs v. Duke Power Co., 401 U.S. 424, 430 n.6 (1971) (comparison of census statistics on percentages of blacks and whites in North Carolina meeting high school diploma requirement and EEOC statistics on pass/fail rates on scored tests by race sufficient to establish prima facie case).

reasons that do not support an inference of discrimination. Although the courts have acknowledged that discrimination can be established entirely on the basis of statistical proof, most cases couple statistical proof with nonstatistical evidence of discrimination. Often statistics are "buttressed" with testimony regarding individual instances of discrimination. The courts, however, have been most impressed when the plaintiff has demonstrated that present procedures perpetuate the effects of past discrimination. Early transfer and seniority cases under Title VII suggest that the courts found currently neutral procedures objectionable when they locked minorities into disadvantaged positions resulting from previous discriminatory treatment by the employer. In hiring cases, the courts have simi-


63. See, e.g., Hazelwood School Dist. v. United States, 433 U.S. 299, 303-06 (1977); International Bhd. of Teamsters v. United States, 431 U.S. 324, 338 (1977); Note, supra note 62, at 472 (many courts require that defendants prove at least one specific instance of discrimination in order to prevail).

64. See, e.g., Local 189, United Papermakers & Paperworkers v. United States, 416 F.2d 980 (5th Cir. 1969), cert. denied, 397 U.S. 919 (1970); Quarles v. Philip Morris, Inc., 279 F. Supp. 505 (E.D. Va. 1968). When transfer or seniority arrangements have locked minorities into less desirable jobs, courts have held that the employer must meet a high standard to prove the business necessity of the practice. See, e.g., United States v. Bethlehem Steel Corp., 446 F.2d 652, 662 (2d Cir. 1971) (necessity connotes irresistible demand); Robinson v. Lorillard Corp., 444 F.2d 791, 798 (4th Cir.), cert. dismissed, 404 U.S. 1006 (1971) (overriding legitimate business purpose required). The Supreme Court has, however, been reluctant to impose a full-scale reorganization of transfer and seniority rights, even when past discrimination is being perpetuated, because of the implications for the expectation interests of incumbent workers. See International Bhd. of Teamsters v. United States, 431 U.S. 324, 353 (1977). Teamsters relied on § 703(h) of Title VII, 42 U.S.C. § 2000e-2(h) (1976), which provides that it is not unlawful for an employer to apply "different standards . . . pursuant to a bona fide seniority . . . system . . . provided that such differences are not the result of an intention to discriminate," 431 U.S. at 345 (emphasis added). In keeping with the concerns expressed in Teamsters, the Court has limited litigant's access to remedies for the effects of transfer and seniority policies. See United Air Lines v. Evans, 431 U.S. 553 (1977) (seniority system giving present effect to past discrimination must be linked to present violation; plaintiff's failure to file timely charge regarding discriminatory termination that severed her seniority rights precludes successful challenge to seniority system); Franks v. Bowman Transp. Co., 424 U.S. 747 (1976) (effects of discriminatory refusal to hire combined with seniority system that perpetuates disadvantage justifies award of seniority relief, but only to individual class members for whom employer cannot successfully show that refusal to hire was nondiscriminatory). The limita-

1465
larly looked to the employer's decisionmaking procedures and their past effects, and examined whether the employer has made any significant change that could reasonably be expected to avoid the effects of past practices.\textsuperscript{63} A finding of present discrimination may be avoided by the employer, however, if the previous practices have been altered and, as a result, the employer can adduce statistics demonstrating no observable disparate effects arising from present practices.\textsuperscript{64} Regardless of the lack of aggregate impact, the plaintiff may attempt to prove disparate treatment by showing that screening procedures were discriminatorily applied to each individual.\textsuperscript{65}

The prima facie showing of disparate impact has been used by the courts in the employment context to evaluate different kinds of decisionmaking processes. In cases of unstructured subjective screening involving ambiguous and undisclosed standards and possibly biased individual judgments, the courts have emphasized the past and present results of the process, the extent of the disparate results, and the employer's lack of effort to improve procedures.\textsuperscript{66} When objective measures are challenged, the courts can more easily evaluate the standards used, in isolation from the overall results of the procedures.\textsuperscript{67} The courts have emphasized the relationship between

\textsuperscript{63} Hazelwood School Dist. v. United States, 433 U.S. 299, 301-04 (1977) (unstructured screening process with historic discriminatory effect and efforts to recruit teachers only at predominantly white educational institutions continued unmodified by employer).

\textsuperscript{64} See Furnco Constr. Corp. v. Waters, 438 U.S. 567, 571-72 (1978) (case on disparate-impact theory avoided by employer with statistics demonstrating that labor-force parity was achieved by affirmative recruitment voluntarily instituted). An employer cannot, however, by altering his present practices, erase his liability for past acts of discrimination. Cf. id. at 584 (Marshall, J., dissenting) (“employer cannot be relieved of responsibility for past discriminatory practices merely by undertaking affirmative action”).

\textsuperscript{65} See id. at 579 (“A racially balanced workforce cannot immunize an employer from liability for specific acts of discrimination.”); International Bhd. of Teamsters v. United States, 431 U.S. 324, 334-43 (1977) (developing distinction between disparate treatment of individual and disparate impact on class of individuals); cf. McDonnell Douglas Corp. v. Green, 411 U.S. 792, 802 (1972) (complainant establishes individual prima facie case by showing (i) individual belongs to racial minority; (ii) he applied and was qualified for job for which employer was seeking applicants; (iii) he was rejected; and (iv) position remained open).

\textsuperscript{66} See, e.g., Furnco Constr. Corp. v. Waters, 438 U.S. 567, 570 (1978) (deliberate effort by employer to identify and recruit qualified blacks); Hazelwood School Dist. v. United States, 433 U.S. 299, 301-04 (1977) (failure to improve unstructured hiring standards and procedures and failure to recruit at minority colleges, although actively recruiting at predominantly white institutions).

\textsuperscript{67} The problems inherent in evaluating standards separately from their overall result were particularly salient in Washington v. Davis, 426 U.S. 229 (1976) (applying equal protection clause) (verbal-ability test used to screen job applicants). In Washington v. Davis, plaintiffs produced evidence that the verbal-ability test used as a hiring standard
these standards and the tasks to be performed, and have disallowed broad standards that would automatically exclude minorities when a more particularized standard might not. 70 Because it is often difficult to assess the relationship of scored tests to the tasks for which they screen, regulatory agencies have set out methodological guidelines for the fair use of such tests 71 and the courts have relied on these standards. 72

2. Business Necessity

If a prima facie case is made by the plaintiff, the defendant may still rebut it by showing that the practice or standard is necessary for police officers led to rejection of blacks over whites at a four-to-one rate. They attempted to prove that the relationship between the verbal-test score and job performance was not established. Defendants, however, had demonstrated a vigorous program of black recruitment and produced undisputed evidence that the result of the hiring process was substantially similar to population percentages in the region. Further, defendants showed that success on the verbal-ability test correlated with success in training school. One commentator argued that the disparate rejection rate for minorities in Washington v. Davis was a concomitant of the vigorous recruitment of black applicants and that a fairer standard for evaluating the test and screening procedure, at least for complex jobs, would be comparison of percentages in a qualified pool in the region, rather than comparison with an applicant pool. See Lerner, supra note 60, at 269-79. The Supreme Court appears implicitly to have acknowledged this reasoning. See p. 1464 supra. A policy of encouraging good-faith recruitment efforts would seem to support application of the Washington v. Davis approach in statutory cases as well.

Still, one can imagine a circumstance under which recruitment practices would not adequately explain the disparate rejection rate: an employer, although seeking majority applicants of average qualification, might recruit only highly qualified minorities to achieve general population parity and employ a screening device that was not well related to job needs and was highly discriminatory against minorities generally. Cf. Note, supra note 62, at 470 (unless validated as job-related, general intelligence test standard accepting whites at 10-times rate of blacks, in region where 10% of population is black, is discriminatory, if 50% of available union members are black).

70. Rejecting the employer's use of high school diploma requirements and objective scored tests in Griggs v. Duke Power Co., 401 U.S. 424 (1971), the Court highlighted the goal of ensuring that screening standards allow individuals an equal opportunity to show job-worthiness. The objectionable features of the employer's standards were threefold. They were overbroad and general. Id. at 433. Because they included attributes that minorities had not had an equal opportunity to attain, the standards tended to "freeze" minorities into lower-paying positions. Id. at 430. Moreover, the standards did not allow the basic intelligence of a minority-group member "the means of articulation to manifest itself fairly." Id.

71. See 29 C.F.R. § 1607.1 (1978) (guidelines to employers to determine whether tests are in compliance); 41 C.F.R. § 60-3 (1978) (similar guidelines for government contractors). The American Psychological Association has issued guidelines concerned with such test development. See, e.g., AMERICAN PSYCHOLOGICAL ASSOCIATION, STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTS AND MANUALS (1966) (general standards and guidelines for evaluative information to be provided by professional test developers to user); DIVISION OF INDUSTRIAL-ORGANIZATIONAL PSYCHOLOGY, AMERICAN PSYCHOLOGICAL ASSOCIATION, PRINCIPLES FOR THE VALIDATION AND USE OF PERSONNEL SELECTION PROCEDURES (1975), reprinted in B. SCHLEI & P. GROSSMAN, supra note 53, at 1395-1412 (guidelines for testing practitioners).

for a legitimate business purpose. The business necessity concept remains vague and necessarily tied to the underlying facts, but some general principles can be delineated. The courts have long acknowledged that the defendant has a legitimate interest in hiring and promoting employees who can perform the particular tasks required. However, in *Griggs v. Duke Power Co.*, the Supreme Court held that a standard allegedly designed to improve the overall quality of the work force was unacceptable because it did not have a “manifest relationship to the employment in question.” It emphasized that the social preference embodied in discrimination law is for standards and screening procedures that allow individuals, regardless of group membership, an opportunity to demonstrate job-worthiness. Recognizing that use of individual screening or validated tests to evaluate job capabilities may be more costly than broad standard requirements or unvalidated tests, the courts have indicated, nonetheless, that the legitimate desire for inexpensive and rapid evaluation procedures and standards does not constitute a business necessity defense for a test or standard that is not adequately job-related.

73. 401 U.S. 424 (1971).
74. The Court took as probative the fact that “employees who have not completed high school or taken the tests have continued to perform satisfactorily and make progress in departments for which the high school and test criteria are now used.” *Id.* at 431-32. In rejecting the employer’s justification in *Griggs*, the Court equated business purpose with job-relatedness. For a discussion of the conflict between a broad business purpose standard and a job-relatedness standard in the lower courts prior to *Griggs*, see *Developments in the Law—Employment Discrimination and Title VII of the Civil Rights Act of 1964*, 84 HARV. L. REV. 1109, 1132-39 (1971).
75. 401 U.S. at 432. The Court in *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975), amplified these concerns when it addressed standards similar to those in *Griggs*, but supported by somewhat more sophisticated validation. The Court found objectionable the fact that the employer relied on general-ability tests without relating the measures to particular attributes or skills needed in various job groupings. The validation was also weak because it correlated scored tests with job performance as measured only by general supervisory rankings that were based on job groups near the top of lines of progression. Finally, the validation study was based on the performance of white workers only. *Id.* at 431-35.
76. *See* note 70 *infra*.
77. *See*, e.g., *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 427-36, 428 n.24 (1975) (general-ability tests rejected because they tested broader range of skills than job requirements justified and local validation of tests was inadequate); *Robinson v. Lorillard Corp.*, 444 F.2d 791, 798-99 (4th Cir.), cert. dismissed, 404 U.S. 1006 (1971) (industry practice, efficiency of promoting those with most experience to maintain employee morale, and avoidance of union strikes over contract provisions, although related to business costs, do not show business necessity justifying discriminatory departmental seniority system). The *Robinson* court noted: “While considerations of economy and efficiency will often be relevant to determining the existence of business necessity, dollar cost alone is not determinative. For example, although there undoubtedly are significant costs involved in validating tests, *Griggs* requires that employment tests be abandoned if not specifically validated as job-related.” *Id.* at 799 n.8; see *Johnson v. Pike Corp.*, 332 F. Supp. 490, 495-96 (C.D. Cal. 1971) (garnishment of employee’s wages as basis for discharge) (“only
Credit Scoring

In considering business necessity justifications, the courts have shown more concern about threats to life than about threats to profits. In early seniority and transfer cases involving racial discrimination, business necessity was defined as a practice that is necessary for the "safe and efficient" operation of the business. A similar requirement evolved as the basis of the defense that sex was a "bona fide occupational qualification." When efficiency is at issue, the courts have generally scrutinized alleged business necessity defenses closely. In contrast, a threat to the safety and well-being of others has justified the use of standards that are exceptionally broad or are applied in a nonparticularized way.

Two questions regarding the standard for business necessity remain open. The first is the extent to which an enterprise must particularize and validate its measures according to specific performance categories. The second and more important question is when efficiency or cost concerns are legitimate defenses.


79. Proof that sex, national origin, or religion is a "bona fide occupational qualification" is a defense under Title VII for the consideration of these characteristics in an employment decision. 42 U.S.C. § 2000e-2(e) (1976); see, e.g., Diaz v. Pan Am. World Airways, 422 F.2d 385, 388 (5th Cir.), cert. denied, 404 U.S. 950 (1971) ("business necessity" not "business convenience" test; discrimination on basis of sex valid only when essence of business operation would be undermined by not hiring one sex exclusively); Weeks v. Southern Bell Tel. & Tel. Co., 408 F.2d 228, 233 (5th Cir. 1969) (to rely on bona fide occupational qualification exception, employer has burden of proving that all or substantially all women cannot perform safely and efficiently).

80. See Dothard v. Rawlinson, 433 U.S. 321, 333-36 (1977) (bona fide occupational qualification grounds) (essence of correctional counselor's job is to maintain prison security; at least under conditions in instant case, presence of woman would pose real threat to security); Spurlock v. United Airlines, 475 F.2d 216, 219 (10th Cir. 1972) (disparate impact grounds) ("When the job clearly requires a high degree of skill and the economic and human risks involved in hiring an unqualified applicant are great, the employer bears a correspondingly lighter burden to show that his employment criteria are job-related.")

81. Cost concerns have received some deference from the Court in the recent pregnancy-benefits cases. See Nashville Gas Co. v. Satty, 434 U.S. 136 (1977) (although employer may not suspend seniority rights for women taking pregnancy leave, he could refuse to grant disability pay at least when no showing that men receive more benefits than women); General Elec. Co. v. Gilbert, 429 U.S. 125, 136-39 (1976) (employer's plan excluding pregnancy benefits not sex discrimination per se because it included no risk from which only one sex protected; no disparate impact when pregnancy benefits costly and benefits package lacking pregnancy benefits was not worth more to men than to women). It should be noted that these cases have been legislatively amended. Act of Oct. 31, 1978, Pub. L. No. 95-555, 92 Stat. 2076 (amending 42 U.S.C. § 2000e (1964)) (adding § 701(k) of
B. The Analogy Between Credit Screening and Employment

1. Similarities

The granting of credit is similar to hiring in several ways. Both decisions involve assessments of risk. The primary problem for the screening party in each case is minimizing the risk of bad performance and loss associated with a relationship yet to be created. The problem for the individual applicant is signaling his qualifications to the screening party.

Similarly, the use of statistical screening devices in credit granting resembles the hiring process. The devices serve the same basic purposes in both areas: they are intended to increase the screener's predictive power while avoiding the expense of making individualized determinations. They also create the same types of dangers in both areas. As has already been discussed, a statistical device can discriminate explicitly, by using overbroad or exclusive measures, or implicitly, by being based on prior exclusionary practices or by incorporating stereotyped or otherwise inaccurate group characterizations. As a result, the devices can underpredict or mispredict for protected classes.

The similarity between credit granting and employment suggests Title VII to clarify that denial of pregnancy-related benefits constitutes sex discrimination.

Sex-related cost differences did not justify discrimination in Los Angeles Water & Power Dep't v. Manhart, 435 U.S. 702 (1978), in which the Court concluded that the Title VII requirement of sex-neutral treatment forbids assessing pension plan contributions on the basis of a sex-differentiated mortality table. It rejected the argument that differential treatment was excused by a "difference in the cost of providing pension benefits," and stated simply that Title VII does not contain a "cost-justification defense." Id. at 716.

The Court did, however, conclude that difficult conceptual problems of fairness, and social policy considerations associated with benefits estimation precluded awarding back pay to females who had previously contributed to the plan at a higher rate than males. Id. at 719-23. In so doing, the Court seemed to be taking account of both cost considerations and the egregiousness of the wrongdoing at issue.


83. The relationship between the screening party and the applicant, although reciprocal, is not symmetrical. The screening party has a long-term interest in the signaling exchange and can choose signals to suit himself. In contrast, the applicant may have little power, in the short run, to make a favorable showing. Id. at 74-75 (in credit market, many potential signals (arguably alterable characteristics) turn into effective indices (unalterable characteristics) as costs of altering characteristics exceed gains for individual applicant). Furthermore, it is unlikely that the individual applicant has the time, resources, or interest to obtain an understanding of the signals used or to develop a full repertoire of meaningful signals. Id. at 107 (relatively large number of people in market at one time and relative infrequency with which one individual appears in market conspire to eliminate investment in signaling by primary signalers).

84. See pp. 1455-58 supra.
Credit Scoring

that the basic outlines of the effects test as applied to employment can also be applied to credit granting. In both cases, a statistically significant difference in acceptance rates may be taken to establish a prima facie case of discrimination. The burden of remedying this situation is correctly placed on the screener because he alone possesses the means to reorganize screening and provide equal access. Especially when the screener looks to attributes that the individual applicant cannot control, or that are often not directly related to actual individual performance, the burden of changing the screening device should be placed on the screening party. In these respects, the creditor may differ from the employer because the creditor more frequently relies on such unalterable attributes. Therefore, the burden of remedying the screening defects should fall more heavily on the creditor than on the employer.

There is a final similarity between the granting of credit and employment. Both are socially valued activities. The practical need to judge people according to their work skill or their creditworthiness must not be forgotten as the law attempts to remedy discrimination. The social goals associated with credit and employment are important to the public at large and they are shared by groups protected by the ECOA. Employment and credit opportunities are made possible by business health and prosperity. Little would be gained by society if courts required access for those who were not qualified. Thus, a business necessity defense in both areas makes sense, despite the shortrun disadvantages imposed on protected classes.

2. Differences Between Credit Screening and Employment

In framing the judicial test for discrimination in the credit context, it is also important to note some differences between credit screening and similar screening in the employment context. Credit

85. Cf. G. Calabresi, The Costs of Accidents 68-94 (1970) (allocation of risk in tort law). In addition to being the primary party with the means to recognize screening, the creditor may be the most efficient means of distributing the costs of eliminating discrimination. The cost of reorganizing credit screening can be distributed onto society as a whole through other avenues as well. For example, government programs offer direct loans and credit insurance to remove or mitigate imperfections that keep the lending market from delivering funds to equal risks at equal cost and to encourage economic transactions, such as home ownership and education, that benefit society as well as the purchaser. D. Larrin, $300 Billion in Loans: An Introduction to Federal Credit Programs 4-5 (1972). However, because the screening party has a long-term interest in the screening process, and the capacity to determine which signals to recognize, see generally, A. Spence, supra note 82, the screening party is a readily available avenue of redistribution and has incentive to find the most efficient procedures that meet standards of fairness.

86. See note 85 supra; pp. 1472-73 infra.
scoring may have more objectionable features than employment testing. In addition, the use of objective predictors in the screening process is relatively more important for credit granting than for employment.

First, problems in the credit area are not as susceptible to easy quantitative solutions. Creditors are not allowed, under the Act, to validate the impact of a credit-scoring scheme by separating credit recipients according to membership in protected groups. In addition, it is much more difficult to particularize the criteria tested in the credit context than in hiring because of the broad group characteristics necessarily used in the actuarial-like prediction of credit scoring and because there seem to be many kinds of personal attributes that are relevant to creditworthiness. Thus there is a need

87. Regulation B and regulatory agencies generally forbid creditors to collect information on individual applicants that would allow them to assess the differential validity of the scoring instrument on the creditor's own pool. 12 C.F.R. § 202.5(d) (1978). But see id. § 202.13 (creditor who receives consumer applications for purchase of residential real property shall request information regarding protected status for monitoring purposes). Some regulators and consumer advocates opposed proposals to allow creditors to inquire about protected group membership because, they assert, such inquiries would facilitate willful discrimination and make the task of proving such discrimination more difficult.

88. Although credit scoring resembles actuarial prediction in that it employs broad groups to predict risk, credit scoring differs as a practical matter in at least two important respects. First, credit scoring is used to predict an event—nonpayment—that is likely to occur with only a small subset of borrowers whereas actuarial prediction for life insurance and annuities predicts the timing of an event—death—that occurs with certainty. Second, credit scoring is almost always based on a creditor's own sample of past borrowers, while actuarial prediction is often based on industry-wide data. These differences, along with others, mean that credit scoring is generally less sophisticated and less statistically reliable than actuarial prediction for insurance purposes. See 1975 House Hearings, supra note 45, at 37 (statement of Jeffrey M. Bucher, member, Federal Reserve Board).

89. See pp. 1453-54 supra. Typically a job applicant is screened primarily for relevant job skills. Employment tests and other standards are devised to measure in some rough manner the extent of such skills possessed by the individual applicant. By contrast, credit scoring does not attempt to assess the level of individual achievement or performance history but only to assess whatever small probability of nonpayment exists on the basis of broad group categories. Credit payment performance could be and is used by some creditors as a screening device and good payment history, once established, may be the best predictor that the individual will show future good payment. Cf. Stanford Research Institute, supra note 21, at 98 (analysis suggests that, even for quite new accounts, customer's record of charging and payments is of so much importance as to almost entirely outweigh value of data contained in original application). However, the costs and time required in obtaining credit reports, and various shortcomings of credit reporting practices, have discouraged some creditors from placing primary reliance on credit reports for new customer screening. Further, credit history standards disadvantage women and minorities who have been excluded from receipt of credit. See note 103 infra. Screening with tests that assess individual attitudes based on a behavioral model has been proposed. See Grablowsky, Behavioral Model of Risk in Consumer Credit, 30 J. Finance 915 (1975). But this approach has been considered impractical because fraud is easier with such a screening device than with prediction devices keyed to independently verifiable information. See B. Shinkel, Effects of Limiting Information in the Granting of Consumer Credit 41-42 (1976) (unpublished Ph.D. dissertation) (on file with Yale Law Journal).
Credit Scoring

for different and more creative judicial and administrative solutions in the credit context.

A related difference is that creditors rely more often than do employers on predictive variables that individuals cannot or typically will not change for the purpose of getting credit. Although this problem is present to some extent in employment testing, the recent efforts to relate measures to particular tasks have made it easier for protected groups to perform favorably on scored tests. Efforts to improve the availability of education and training represent an additional mitigating factor. The nature of credit prediction generally suggests that it is unreasonable to rely on similar avenues to improve credit access for creditworthy members of protected classes, and that the bulk of the accessibility problems must be confronted by altering the organization and standards of screening itself. Consequently, creditors should have a greater responsibility than employers to discover and overcome the differential effect that their chosen measures have on creditworthy members of protected classes.

A second major difference between the use of scored tests in employment and credit screening lies in the relative importance of risk prediction to the success of the enterprises involved. The employer's ability to identify all the job applicants who represent the best risk has only a limited relationship to profitmaking. His desire and capacity to hire new employees are only partially dependent on his ability to locate potential good employees. Although discrimination laws may force an employer to search out new sources of able employees among protected classes, these laws cannot, by doing so, significantly increase an employer's profits or expand the total number of jobs available.

Creditors have a greater interest in developing effective predictive-screening devices because credit screening is an aspect of marketing and sales. The number of credit applicants that a creditor may ac-

90. See p. 1454 supra. In order to decrease information cost and achieve reliability of prediction, creditors prefer to use characteristics that individuals cannot alter or generally do not change to acquire credit. Cf. A. Spence, supra note 82, at 74 (costs to individual applicants of altering signals exceed gains); B. Shinkel, supra note 89, at 41-42 (risks of "gaming" with credit-screening devices employing information easily altered or not readily verified by creditor).

91. See p. 1468 supra.

92. During the last decade, one alternative to requiring employers to hire untrained or relatively less skilled employees to overcome the effects of past exclusion has been to provide job-training programs for disadvantaged groups. See C. Bell, The Economics of the Ghetto 174-93 (1970). Although the costs of such an approach may have been great, the strategy is in keeping with the general societal value placed on training and experience.

93. See p. 1471 & note 83 supra.
cept is a function of expected risk, costs, and the number of creditworthy persons willing to purchase credit from that creditor. The creditor seeks to improve the predictive accuracy of his screening in order to increase his ability to borrow or resell his consumer portfolio to the secondary lending community. Further, unlike an employer, the creditor himself will undoubtedly benefit from expanding his applicant pool. A creditor who can identify previously excluded individuals who in fact are good credit risks can expand his operations and his profitability. Therefore, requiring creditors to alter screening practices to identify new good risks is less burdensome, in the long run, than requiring employers to make comparable alterations.

III. A Proposed Standard for the Effects Test in Credit Screening

On the basis of the foregoing analysis of the effects test, it is possible to propose a standard for credit scoring that deals with the conflict between equality and predictive convenience. To fulfill the goal of equal opportunity, screening must allow individuals who would perform equally well an equal chance to signal their expected performance. Thus, the effects test should be applied in a manner analogous to its application in the employment context. Because the creditor himself would benefit from the increased applicant pool that would emerge under a restructured screening device, this standard should be more stringent. In cases in which the standard is not met, creditors should be required to take affirmative steps to implement screening techniques that improve accessibility for minorities and women.

Existing discrimination law is notably unspecific about an "affirmative steps" requirement because it entails a case-by-case balancing of competing interests. In the credit-scoring context, however, the pri-
mary emphasis should be on preventing discrimination because the overall benefits of removing discrimination outweigh countervailing concerns. Fairness to creditors does not require leniency, but it does require advance notice of the standard that they must meet.

A. The Prima Facie Case

The standard for proof of the prima facie case should implement equal-access goals while enabling the creditor to know what screening practices may be discriminatory. Formulation of such a standard involves two major determinations—how the relevant population for comparative analysis will be defined, and how stringent the requirements of equality will be.

In discrimination law the accepted universe of comparison is the applicant pool or qualified population in the relevant geographic region. For credit discrimination, the qualified population should include all persons who fall within the general range of a reasonable minimum income and do not have incomes exceeding a certain maximum associated with the lender’s risk level. Minimum income can, in turn, be defined by the amount of income necessary to meet the usual obligation to the creditor. In determining this minimum, the

1977, cert. granted, 99 S. Ct. 608 (1978) (No. 78-435) (voluntary affirmative job-training quota for racial minorities impermissible when quota not instituted to correct past discriminatory practices). Note that Title VII includes an “antipreferential” provision, 42 U.S.C. § 2000e-2(j) (1976); except insofar as provided by the general requirement not to discriminate, the ECOA contains no similar provisions. Although H.R. 6516, one of proposed amendments in 1976, contained a provision, § 703(b), which stated that failure of a creditor to achieve minority representation proportional to the general population was not a per se violation of the Act, H.R. 6516, 94th Cong., 1st Sess. (1976). That provision was omitted in the enacted amendment.

98. See pp. 1463-64 supra. The relevant geographic region has never been reduced to a uniform standard. When dealing with pass-fail comparisons, the courts have considered the geographic area from which applicants are likely to come. For general population-work-force comparisons, the courts have applied a more restrictive standard. See B. SCHLEI & P. GROSSMAN, supra note 53, at 1178-81. The geographic scope of comparison must similarly be determined on a case-by-case basis for creditors.

However, because minorities in urban areas tend to live in predominantly minority residential neighborhoods, see M. DANIELSON, supra note 31, at 10-11, and higher rejection rates for minority credit applicants are observed even within income and asset categories, see 1975 Senate Hearings, supra note 33, at 461-62 (statement of Steven M. Rohde); COMPTROLLER OF CURRENCY, supra note 33, minorities have been forced to pay more for goods, D. CAPLOVITZ, THE POOR PAY MORE 49-93 (1967), and credit, id. at 94-104. Thus, a minority neighborhood should not be excluded from the scope of comparison because inhabitants of those neighborhoods have relatively low incomes or do not show a high rate of trade with the creditor.

99. The creditor’s primary concern, ability to repay, would generally be met by the income standard. Persons who met this standard would also be likely to have assets or possessions that could be recovered in the event of default. But see note 102 infra.

Analyzing the creditor’s past income standards and the general income range of the individuals in his credit pool serves two functions. First, when the availability of a sizeable down payment is a factor in the loan, the creditor’s experience suggests an income
cost of living in the region and the creditor's experience with income levels that have high default rates is relevant. The income range represented in a creditor's historical pool of debtors would also be probative of fair criteria for the population of comparison, especially if the creditor has not recently raised the floor.100

As an aggregate measure, the income-range test meets the general statutory standard of willingness and ability to repay.101 This same standard provides a fair percentage of the population for comparison in the judicial test;102 it is neither unreasonable to the creditor range of individuals who have met the standard. Second, such an analysis accounts for industry price stratification. See note 18 supra. Industry stratification implies that each creditor deals primarily with a population composed of individuals within a given income range. Because most protected classes constitute a higher proportion of low-income populations than they constitute of the general population, the courts should employ an income ceiling as well as an income floor in applying the effects test, at least for some creditors. Otherwise, the judicial test will protect creditors at the upper end of the income-risk continuum and will be lax with creditors at the lower end, where most protected classes are overrepresented compared to their proportion in the general population. Cf. Note, 1976 Amendments to the Equal Credit Opportunity Act, 28 BAYLOR L. REV. 633, 647 (1976) (unless "adverse action" provisions of ECOA applied in manner that takes account of risk-reflective interest rates, availability of credit from creditors who do adjust credit terms to lend to broad range of population will be restricted).

100. If the rise in minimum income were in keeping with the rate of inflation, this could easily be demonstrated at trial. However, significant changes made soon after the passage of the Act or its amendments would support an inference of discrimination. Under the effects test, the creditor would have the burden of proving that the change had a bona fide purpose if the change contributed to a statistically significant disparity.

101. See note 44 supra.

102. The ability to pay reflected in income is, of course, mediated by an individual's savings rate. The Act disfavors an assumption that members of one group will have higher obligations or lower savings rates than those of another; however, there may be cases in which this is true. For example, members of one group, for reasons of cultural preference, may tend to have more dependents, and therefore lower economic capability at a given income level. Of course, the creditor is prohibited from using group membership to predict this relationship; he may, however, use number of dependents as a predictor of payment. The effects test allows creditor screening by multifactor credit scoring to defeat the prima facie case upon a showing that: (a) the number of dependents is, in fact, related to payment behavior for both groups and (b) the disparate impact of the scoring device that contains several variables can be substantially explained by this variable. See Note, supra note 61. In contrast, a defendant need not make the second showing when the criterion relied on is a unitary pass-fail device and the permissibility of any component factor is not an issue.

The creditor who had previously excluded one group would have difficulty demonstrating statistically that the number of dependents related to payment behavior for the group. This demonstration is important because the effect of the number of dependents on credit payment ability is probably greater at the low end of the income continuum than at the high end. If the group with whom number of dependents was derived as a predictive variable tended to have more dependents among its lower-income members while the new group tended to have more dependents at the upper-income level, a low score for high number of dependents might be overburdensome for the new group.

The court might look to other evidence in order to determine whether the relationship of number of dependents to other factors, such as income, in the two groups is similar enough that the variable would be expected to operate similarly as a predictor for both groups.
Credit Scoring

nor so lax that it merely reflects past discriminatory practices. An income-based comparison pool for the effects test would not, of course, imply that a creditor must grant credit on the basis of income. He may find that other variables are better predictors of repayment for him. The income criterion is merely used to alert the court to instances when a creditor may have failed to take sufficient steps to mitigate the exclusionary effects of traditional credit-screening practices.

Some commentators have suggested limiting the comparison population to the applicant pool. But such a rule would not force creditors to take affirmative steps to improve the access afforded

103. Financial and credit-reference measures should not be standards for inclusion because they build in the effects of past exclusionary practices of the credit and financial industry. Even as a basis for exclusion, credit references have a number of pitfalls. First, many women and minorities have never received credit and consequently do not have a credit history. See Letter from Lewis H. Goldfarb, assistant director, Sally Gold and Rena Steinzor, attorneys, Division of Credit Practices, Bureau of Consumer Protection, Federal Trade Commission to Theodore E. Allison, secretary, Federal Reserve Board (July 28, 1978) (on file with Yale Law Journal) [hereinafter cited as FTC Letter]. Second, creditors who lend to the lower socioeconomic groups tend to employ more stringent monitoring and credit-reporting techniques so that it is likely that such groups will accumulate a history of late payment. Cf. Stanford Research Institute, supra note 21, at 98 (extension of credit is often not only end in itself but also means to build up proven customers; in general only consumer-finance companies provide “saturation level” collection effort). The court should be cautious in either anchoring proof of discrimination or fashioning remedial plans on the basis of late payment history because such screening may be an artifact of creditor-reporting patterns that vary according to the social class of a creditor’s usual customers.

104. Research on payment prediction suggests that once a threshold level of income is met, income does not operate to predict bad payment. See, e.g., D. Durand, supra note 22, at 4-5, 45-48 (none of income distributions of borrowers from creditors including commercial banks, retail merchants, and finance companies show more than moderate tendency for higher incomes to be better risks). Explanations for this phenomenon may be that stability of income, the margin between income and expenses available to repay indebtedness, and the character and financial awareness of the credit user are more important than his absolute amount of income given a minimum capacity. Id. at 5.

105. See Hsia, supra note 23, at 421 (arguing that applicant pool comparison is preferable alternative because it fits creditor’s unique market). The applicant pool, moreover, is a readily available alternative because Regulation B requires the creditor to retain records on applications and actions taken for 25 months after the applicant receives notification from the creditor. 12 C.F.R. § 202.12(b) (1978). But cf. Note, supra note 62, at 409-72 (illustrating various conditions under which comparative (pass-fail) and demographic (general population) statistics might be beneficially used in evaluating employment discrimination).

106. In fact, use of the applicant pool might discourage affirmative recruitment. Cf. Lerner, supra note 60, at 272. Furthermore, a creditor who did not have significant experience with protected classes might avoid these groups in order to avoid consideration of new kinds of applicants for which his past sample was not an adequate representative. Cf. D. Durand, supra note 22, at 7 (concern in credit prediction that scoring necessarily derives from sample of past credit users); Eisenbeis, supra note 25 (problems of measurement in credit prediction include sample bias); FTC Letter, supra note 103 (discussing creditor’s techniques for avoiding dealing with general population). Because of the stratification of the market, see note 18 supra, the court may want to consider both the applicant pool and the relevant population in the region to assess the fairness of the creditor’s standards and procedures.
by screening. In addition, applicant-pool comparison is less likely to alert courts that creditors are avoiding or discouraging applications from certain subpopulations.\(^\text{107}\) Although such avoidance or discouragement might be treated as another form of discrimination,\(^\text{108}\) proof problems might be insurmountable. Prescreening and solicitation techniques are not easily detectable by private parties unfamiliar with credit practices.\(^\text{109}\)

Once the relevant population has been defined, the final question concerns the degree of stringency appropriate in applying the nondiscrimination requirement. The nature of consumer credit and the types of measurement used both suggest the adoption of demanding requirements. A significant statistical difference in credit access between qualified members of a protected class and members of the majority group should be sufficient to establish a prima facie case.

Creditors will argue that such a standard is overburdensome and,

107. Cf. FTC Letter, supra note 103. In this letter, the staff of the FTC recommended to the Federal Reserve Board that Regulation B be amended to define "credit transaction" to include direct mail or telephone solicitation of prospective applicants since many creditors use criteria and procedures to select prescreened lists that have the "effect of excluding a disproportionate share of women and minorities within the market area." Id. The Commission noted that lists are often screened by zip code to isolate individuals with desirable demographic characteristics. Id. (citing REPORT OF THE PRIVACY PROTECTION STUDY COMMISSION, PERSONAL PRIVACY IN AN INFORMATION SOCIETY 136-37 (1977)). The FTC Letter also discussed a letter from a national retail-credit bureau to a large national credit-card issuer that outlined the criteria used for a prescreened solicitation. The criteria included:

1. Minimum file history of two years. 2. Minimum of three satisfactory trade items rated open/paid satisfactorily from the following subscriber industries: bank, bank credit card, retail credit card, sales finance, credit union, savings and loan, service and professional. 3. Maximum of three inquiries within six months. 4. Absence of any trade item from the loan finance industry. 5. Absence of any Special Comments or Manager Attention transactions.

FTC Letter, supra note 103. The FTC recited the manner in which each criterion operates to exclude on prohibited bases. Id.

108. See Note, supra note 8, at 910-11 (suggesting that proof of employer practices discouraging certain groups from taking test should be basis for prima facie case of discrimination).

Regulation B, 12 C.F.R. § 202.5(a) (1978), specifically prohibits the creditor from making any oral or written statement that would discourage members of protected classes from applying. It does not, however, proscribe as discriminatory a general solicitation strategy that simply avoids protected classes.

109. The possibility that the creditor will solicit using mailing lists, with unidentified demographic characteristics, from an independent source or apply his own evaluative standards only after they have been filtered through another source, see note 107 supra, poses elusive, difficult and costly discovery problems for the private litigant in addition to those of garnering documents from the creditor, cf. Hsia, supra note 23, at 422 (obstacles to obtaining creditor data on credit applicants includes privacy rights of applicants, protective restrictions of Fair Credit Reporting Act on disclosure of applicant information and proprietary nature of material), of conducting computer analysis of credit scoring, of developing statistical proof, and of certifying class actions.
Credit Scoring

therefore, counterproductive. In supporting their argument they may invoke several facts that seem initially persuasive. First, unless the scoring system is a poor predictor for the overrepresented group as well, the result of modifying a partially effective system could be a worse overall performance of the enterprise, without improved access for protected groups. Further, creditors may argue that the task of devising a scoring system that applies with equal accuracy to all groups is costly and virtually impossible.

Although these concerns are real, they are not sufficiently compelling to prevail. The range of options available to the creditor under the effects test casts doubt on the legitimacy of unmodified credit practices that have a disparate impact. The creditor's legitimate concern is that he not be required to assume bad risks in credit granting simply to achieve parity with the percentage of the population represented by protected groups. The effects-test standard proposed in this Note satisfies this concern. It requires the creditor to achieve parity only within the qualified population, which is defined by income.

Although income is the means proposed here to identify the qualified population of comparison for the prima facie case, such a technique would not require creditors to screen on the basis of an income criterion or on any other particular basis. Although the creditor's screening standard must relate to performance, absent a showing of disparate impact the ECOA does not prefer one particular screening standard or procedure to another, and does not prescribe that a

---

110. See 1975 House Hearings, supra note 45, at 94 (statement of Richard Cremer, Montgomery Ward) (eliminating age as score category would substantially increase losses); Stanford Research Institute, supra note 21, at 101 (credit scoring actually achieved material reductions in loan losses typically in the 20-40% range); B. Shinkel, supra note 89, at 133-65 (impact of removing sensitive variables, such as race and marital status, from models based on finance-company data, may be—depending on the creditor's risk cutoff level—some reduction of percentage of good loans predicted, increase of bad loans accepted, and consequently some loss of profitability).

111. Cf. 1975 House Hearings, supra note 45, at 24 (statement of Jeffrey M. Bucher, member, Federal Reserve Board) (cost of credit will also be increased to extent that alternative tests are more expensive to apply).

112. See Hsia, supra note 23, at 426 (one system may select creditworthy applicants from one group while performing poorly for another; selection may require creditor to make tradeoffs among protected classes); Chandler & Ewert, supra note 25 (illustrating that differential measurement of male and female performance is means to identify most good risks among females, especially without also increasing number of bad male risks accepted).

113. See Furnco Constr. Corp. v. Waters, 438 U.S. 567, 578 (1978) (Title VII) ("Courts are generally less competent than employers to restructure business practices and unless mandated by Congress they should not attempt to do so.")
unitary standard be more or less preferred than a multidimensional one.

It is true that certain predictors are more likely to be objectionable than others. A creditor cannot take prohibited bases into account, except for age and marital status under limited circumstances. Predictors, such as address and credit history, are unusually likely to have a disparate impact on a protected class. A creditor suspecting that his scoring scheme will have a disparate impact may preserve its general form if he also adds screening steps designed to mitigate the expected disparate effects. Using the means available to evaluate the effects of screening criteria, a creditor may survey specific mea-

114. 12 C.F.R. § 202.6(b)(1) (1978) (“[A] creditor shall not take a prohibited basis into account in any system of evaluating the creditworthiness of applicants.”) But see 15 U.S.C. § 1691(a)(3) (1976) (not violation to refuse to extend credit offered pursuant to special-purpose credit program to meet special social needs that fulfill standards prescribed in regulations by Federal Reserve Board).

115. The Act allows inquiry into “marital status . . . for the purpose of ascertaining the creditor's rights and remedies applicable to the particular extension of credit and not to discriminate in a determination of credit-worthiness.” 15 U.S.C. § 1691(b)(1) (1976). It allows inquiry about and consideration of age in order to evaluate particular aspects of the applicant's income, credit history, or other pertinent elements in an empirically derived scoring scheme, if the scheme meets certain requirements, or in order to favor elderly applicants. Id. § 1691(b)(2)-(4).

116. See notes 28 & 31 supra.

117. See notes 33 & 103 supra.

118. But see Hsia, supra note 23, at 429 (because of potentially lighter burden of proof for an empirically derived credit system which is demonstrably and statistically sound, users of “hybrid” system may want to bring their system under that rubric). Many creditors who rely primarily on credit scoring employ a discretionary or judgmental override procedure. Cf. id. at 428 (describing general purpose of judgmental override). The purposes of such a procedure are typically either to check the veracity of a subset of applications or to further examine applications in the middle or lower score range. Although such a hybrid process is a less efficient means of processing applications than scoring alone, it is still less costly and time consuming than screening judgmentally or obtaining credit reports for all applications.

119. Although it is not presently reasonable to expect creditors to validate or score differentially for protected classes, see note 87 supra, this does not mean that creditors are entirely without means to assess the differential impact of their screening strategies and to take some steps to avoid this impact. The creditor has several bases of information that he can use in assessing the possibility that his screening devices make access more difficult for creditworthy members of one group than for another. First, the creditor can evaluate the pre-Act composition of his credit-use pool, see 12 C.F.R. § 202.12(a)(1) (1978) (not violation to retain information collected prior to Act), to determine what kind of population he is measuring. The creditor may also have information about group membership of present applicants from credit reports, see id. § 202.12(a)(2) (not violation to have such information if obtained from consumer reporting agency), that he might use to check applicant-flow composition. Second, the creditor can infer from external information that certain of his measures correlate differently with socioeconomic status for protected groups than for majority-group males. Third, the creditor can perform causal analysis on the measurement factors to determine relevant relationships among predictive factors. See generally H. Blalock, Causal Inferences in Nonexperi-
Credit Scoring

sures on which protected groups are likely to receive lower scores than those of majority groups. The creditor may then identify related characteristics that, by their positive relationship to creditworthiness, justify overriding the general statistically derived score and accepting an otherwise low-scoring applicant. When the creditor needs previously unavailable systematic data on group membership to effect a remedial program, or when he is otherwise unable to satisfy the requirements of the Act, he may opt, under special provisions of the Act, to adopt a special purpose credit program. In addition to strategies to mitigate the disparate effects of old scoring systems by systematic override procedures or special purpose programs, the creditor under some circumstances might achieve the result required by law, and at the same time continue to use the scoring scheme, if he can market in a manner that attracts creditworthy members of protected classes. If the creditor avoids disparate impact in this

MENTAL RESEARCH (1964) (describing techniques and limitations of quantitative analysis for causal inference).

Although causal analysis is not presently used in credit-scoring devices, see Hsia, supra note 23, at 383, the results of an independent analysis may aid the creditor in assessing whether the predictive measures on which he relies may be expected to operate in the same manner for groups with different combinations of characteristics than for his predominant sample. However, the foregoing steps will not be as effective as differential analysis in assessing cumulative disadvantage.

120. For example, the creditor, knowing that minorities are disadvantaged by the use of zip code, may look for factors in an application from that zip code region that indicate that an individual is not a bad risk. Such an approach implies that a credit application from a generally minority zip-code region will get greater scrutiny. But if the creditor achieves a nondisparate result, the approach should not be viewed as discriminatory because it imposes no greater costs on the individual applicant. But see note 103 supra. If the creditor has in fact previously admitted protected class members, he may be able to use his own samples to aid in generating guides for decisionmaking; if not, he must do now what at one time he did with majority group males—work on a hunch.

121. If the creditor finds that he is unable to achieve a satisfactory adjustment through traditional methods or that such an adjustment imposes the risk of a reverse discrimination suit, he may adopt a special purpose credit program. The ECOA allows governmental, nonprofit, and profitmaking creditors to extend credit pursuant to a special purpose credit program. 15 U.S.C. § 1691(c) (1976). Special purpose credit programs of profitmaking creditors must meet standards prescribed in regulations. Id. § 1691(c)(3). Regulation B, 12 C.F.R. § 202.8(a)(3) (1978), requires that profitmaking creditors establish and administer the program pursuant to a written plan that identifies the class of persons benefited and sets forth procedures and standards. Under this subsection, a creditor may establish a special purpose program if he shows that the class of persons to which the program applies probably would not receive credit, or would receive credit on less favorable terms than are ordinarily available to other applicants applying to the organization.

For an example of a special purpose program adopted before the passage of the Amendments, see Brouillette, Response Spurs Wells Fargo to Seek Doubling of Eased-Loan Plan for Low-Income Persons, AM. BANKER, June 6, 1975, reprinted in 1975 Senate Hearings, supra note 33, at 678.

122. See notes 66 & 69 supra.
manner, the burden is on a plaintiff to prove individual disparate treatment.

B. Business Necessity

The extent of the business necessity defense will determine what requirements are imposed on creditors who fail to consider the discriminatory effect of their screening procedures or the effect of past exclusionary practices. A valid business necessity defense cannot be based on difficulties of measurement or the partial effectiveness of a screening method. The creditor's concern with delineating a profitable market is recognized by the use of a qualified pool of comparison, and the interests of society and the creditor in efficient and


124. The shifting burdens of proof under the effects test confirm that the general measures for compliance suggested here are in keeping with the law. If no disparate impact is shown, the creditor need not justify his measures, and it remains for the plaintiff to prove that the standards employed were applied discriminatorily to individual cases. See p. 1466 & note 67 supra. However, one question remains: assuming that protected class members perform better as a class than others in the creditor's pool, can a plaintiff challenge the general screening standard by proving that the qualifications imposed on protected classes were higher than those used for the majority group? The answer depends on what source of information is examined to determine whether higher qualifications were imposed. On the one hand, if the plaintiff can show that the variables predicting creditworthiness used by the creditor imposed higher standards on protected class members than on majority-group males, the answer would be affirmative. For example, the plaintiff may show that the protected class historically confronted difficulty in getting access to credit so that achieving credit history equal to that of white males was effectively more costly for the protected class. In such a situation, it would be unfair for the creditor to score the credit histories of all groups equally. If, on the other hand, the only indication that the qualifications imposed on protected classes were higher was that protected class members evinced the same level of achievement on predictive variables but had a better subsequent risk performance, the answer would be negative. Cf. Los Angeles Dep't of Water & Power v. Manhart, 435 U.S. 702, 708-11 (1978) (invalidating pension contribution schedule based on sex differential mortality tables; Title VII's focus on fairness to individuals precludes treating individuals simply as components of groups). The primary reason is that applying an equal-performance test, based on subsequent empirical performance results alone, in order to evaluate the creditor's standards, implies that we should allow the creditor to distinguish individuals according to group membership and to employ differential prediction. Otherwise he has no guidelines to determine whether he is in compliance with the law.

Because differential screening poses so many technical and conceptual problems of fairness, because Regulation B forbids taking protected class membership into account, see note 87 supra, and because unconstrained availability and use of group-membership information would amplify the problems of regulation, the matter should be addressed by the Federal Reserve Board and other regulatory agencies who could monitor the use of differential prediction if, on a case-by-case basis, it were deemed appropriate.
Credit Scoring

cost-effective screening are served because the effects test defers to the screener's judgment so long as he avoids a disparate outcome.

The law should not go further and recognize higher cost per se as a valid defense under the Act, except in the extreme case.\textsuperscript{125} A broad cost-based business necessity defense would frustrate the purposes of the Act. Dealing with many different groups is always more costly than dealing with one homogeneous group because the creditor is forced to consider a wider range of information and varied patterns of characteristics. Further, he must adjust transactions to meet group differences. Because the creditor lacks information necessary to enable him to find the best possible risks in the new market, it is necessarily more risky for him to operate there. A creditor who previously excluded minorities and women from receipt of his credit is forced to assume these costs; they are costs that were recognized when the Act was passed.

Prior to the Act, the costs of nonparticipation and more difficult access were borne solely by qualified members of protected classes. A stringently enforced effects test would change these costs into higher screening costs distributed to all credit users, not just disadvantaged groups. To accomplish this purpose, a creditor must be prevented from using his present scoring device without mitigating its disparate effects. The fact that he cannot screen protected class members as efficiently as he can screen the majority group cannot be a valid excuse, because such a rule would reintroduce discriminatory practices and transfer the cost back to protected persons. In contrast, the purpose of the Act can be effectively implemented by a judicial standard that requires a creditor to achieve parity within a population that by virtue of income range can be assumed to meet the creditor's contractual requirements. The additional costs imposed by such a standard can be borne by the creditor who ultimately will benefit from the expansion of his applicant pool.

One legitimate concern may argue against an overly constrictive business necessity defense. The process of removing unequal access should not create high costs that fall on innocent third parties who are unable to bear them.\textsuperscript{126} In addition, reorganizing the standards

\textsuperscript{125} Cf. Note, The Cost of Growing Old: Business Necessity and the Age Discrimination in Employment Act, 88 Yale L.J. 565, 587-95 (1979) (discharge of older workers based on claim that older workers with seniority impose higher costs because of higher salaries and fringe benefits should be permitted as nondiscriminatory under Age Discrimination in Employment Act only if economically essential, and if less discriminatory alternatives not practical).

\textsuperscript{126} It was argued on several occasions during the House and Senate subcommittee hearings that creditors can and will mitigate or distribute the costs of nondiscrimination
for distributing benefits may undermine the considerable efforts that individual applicants have already made to meet presently used standards. Upon analysis, however, neither of these considerations provides a generally valid basis for modifying the enforcement of the ECOA.\textsuperscript{127} The third parties who will bear costs will be individual credit users, applicants, and investors. The position of these parties ensures that the costs they must bear will be broadly distributed and occur generally in small amounts that are not unduly burdensome to each party. Credit users may experience a rise in interest rates or the institution of more vigorous collection practices.\textsuperscript{128} The individual credit applicant, now competing with protected class members for credit extensions, may suffer some short-term reduced availability of credit.\textsuperscript{129} The investor may experience a small reduction in

through expanded marketing, increased interest rates, or reduction in assumed risk. The inherent conservatism of credit-granting practice works against expanded marketing; consequently, the initial costs must be distributed through interest rates or risk reduction. See, e.g., 1975 Senate Hearings, supra note 33, at 207 (statement of Jeffrey M. Bucher, member, Federal Reserve Board) (regulation may increase costs and reduce available credit); 1975 House Hearings, supra note 45, at 29 (statement of Jeffrey M. Bucher) (result of increased costs to creditor would be increased costs to other borrowers or higher credit standards).

127. Societal reliance on costly and unique indicators of creditworthiness often corresponds with high expected return on a particular kind of transaction. The most difficult cases in discrimination law, see, e.g., Regents of Univ. of Cal. v. Bakke, 438 U.S. 265 (1978) (challenge to medical school admissions program for disadvantaged minority groups), arise after prior discrimination and the high cost of signaling have discouraged minority participation. See generally A. Spence, supra note 82, at 14-30 (education as signal). The egregious circumstances of discrimination are counterbalanced by the high opportunity costs that third-party applicants have already paid for such traditional indicators of worth. Credit granting is not, as a general rule, based on such indicators acquired in large part for the purpose of obtaining credit. See note 92 supra. A possible exception arises in the housing market in which an individual may have purchased a home in part for its investment and resale value. Even here the possibility of a total or substantial loss on the investment caused by a modification in the credit market is not great.

128. Another means to absorb costs by reducing losses is to impose more stringent controls on accounts either by lowering credit limits or by monitoring accounts more closely. See R. Cole, supra note 12, at 392-39, 566-85. Creditors who have the resources to devise and use credit scoring may also have the computer-system capabilities to monitor accounts and predict default by payment patterns. Such predictive monitoring has the virtue of responding to individual behavior patterns rather than relying on assumptions based on group characteristics and could be implemented as an educative as well as a loss-control device.

129. The factors affecting availability of credit are complex and vary according to kind of credit and region, among other factors. See National Commission on Consumer Finance, supra note 18, at 109-49. One study using data from finance companies to test the effect on access for different demographic groups of removing sensitive variables—race, marital status, age, and occupation—from a scoring device yielded results that generally comport with the purposes of the Act. Groups whose access was improved by removing the sensitive variables included blacks, individuals employed in driver, service or production occupations, individuals married less than 10.5 years, and individuals under 30 years of age. Groups whose access was relatively worse with the ex-
Credit Scoring

investment income. But none of these concerns are in themselves compelling enough to outweigh the interests of protected classes, or of society generally, in achieving equal access.

The concern about devaluation of efforts already made by individuals is closely related to the question of quotas. Initial attempts to allow equal access may be tantamount to the imposition of income-based quotas because of the lack of more sophisticated information. However, the concerns that militate against the use of quotas in employment and higher education are seldom present in credit granting. For example, the employer-employee and school-student relationships are largely exclusive; moreover, once such a commitment is made, the unique quality of the relationship determines later opportunities. In contrast, the creditor-debtor relationship is seldom an exclusive one. Thus, the enforcement of discrimination laws is likely to result in broader distribution of credit, rather than in substantial exclusion of additional individuals from the credit market.

A final concern in formulating a business necessity standard is that the social policy of ensuring access to credit for members of the low-income population may be frustrated. Since participation in the economy is increasingly premised on access to credit, ensuring the availability of credit to low-income individuals is an important social concern. A major argument against the Act was that creditors would inevitably raise the income floor for credit eligibility to reduce risks and offset the cost of nondiscrimination, thereby reducing the amount of credit available to the low-income population. This problem can, however, be avoided. Judicial application of the effects test can take into account the possibility that creditors will use inclusion of sensitive variables included nonminorities, individuals who were single, married, or widowed over 10.5 years, and those over 45 years of age. Effects based on sex were not examined. See B. Shinkel, supra note 89, at 171. The groups benefited by exclusion of sensitive variables were among those for whom concern was expressed during the passage of the Act. Those disadvantaged, except for individuals over retirement age, were those who have been historically more likely to receive credit.

130. See note 119 supra.
132. 121 Cong. Rec. 16,740 (1975) (statement of Rep. Annunzio) (“Credit discrimination cannot be dismissed lightly. Credit has a profound effect on the life of virtually every person in this country. Few of us pay cash. We pay for meals, transportation, homes, cars, hospital care, and numerous other everyday necessities through the extension of credit.”).
133. See NATIONAL COMMISSION ON CONSUMER FINANCE, supra note 18, at 156-58, 160.
134. See 1975 House Hearings, supra note 45, at 24 (statement of Jeffrey M. Bucher, member, Federal Reserve Board) (If to meet costs creditor chooses alternative of raising credit standards, he will deprive marginal borrowers, often those in lower-income brackets, of access to credit).
come standards to exclude creditworthy people with low incomes. If such judicial safeguards are not totally effective and the amount of credit available to low-income applicants is reduced, other measures designed to meet the special problems associated with low-income status can be devised. Such problems do not justify continued discrimination against creditworthy individuals in other income categories.

Various policy considerations support the strict standard advocated here. Permitting creditors to avoid dealing with minority populations affords competitive advantages to discriminators and may injure the overall performance of the economy. Further, credit discrimination forces underrepresented group members to segregate or curtail participation in the economy and to seek credit at higher costs. This contributes to the cumulative disadvantage of women and minorities.

Investigations that led to the passage of the Act suggested strongly that underrepresented groups, particularly women, promised to be an increasingly profitable market, but that creditors were nevertheless slow in adopting nondiscriminatory credit policies. This is explained, in part, by the inherent conservatism of the primary and secondary lending industry. It is also explained by the greater cost of providing credit to a market that includes many different socioeconomic groups. Part of this additional cost will be limited to the initial cost of adjusting past practices; however, even after such adjustments are made, some small costs will remain. It is only such costs, which otherwise would be borne solely by members of underrepresented groups, that the effects tests will impose on creditors and thus on credit users generally.

135. See generally D. Caplovitz, supra note 98, at 180.
136. See generally C. Bell, supra note 92, at 299 (burdens of discrimination on poor);
138. Cf. id. at 139-49 (creditor costs, rates, and availability of credit) (National Commission on Consumer Finance recommending adjustment of interest rate to improve competition in market because present legal restrictions on rates have forced creditors to adjust for costs through reducing risk, which has contributed to reduced availability).
139. Because such costs are an artifact of social organization and measurement problems over which members of protected classes have no control and to which they, as individuals, contribute nothing, under the principles of equal access, these costs should be distributed across the whole population.
140. For example, certain factors will always require more costly, individualized consideration. See 15 U.S.C. § 1691(b)(1), (2) (1976); 12 C.F.R. § 202.6(b)(2)-(5) (1978).