

Noel Capon

# Credit Scoring Systems: A Critical Analysis

The currently fashionable credit scoring systems are described and subjected to critical analysis. Public policy issues concerning the use of these systems are discussed.

"Our society has been taught to believe that an individual's creditworthiness is primarily related to their personal credit history. I feel certain that for anyone who has any regard for the concept of individuality, reviewing the credit-scoring systems of some of our major national creditors would be a chilling experience."<sup>1</sup>

**T**HE importance of consumer credit in the U.S. economy has grown markedly through the 20th century. A combination of growth in the supply and form of credit and increased consumer demand has led to an average annually compounded rate of growth in consumer credit outstanding of 7.5% from 1919, the first year for which Federal Reserve figures are available, to the present. This figure is much greater than the average growth rate of GNP for the same period (Board of Governors of the Federal Reserve System 1976a, 1976b, 1980).

The ever-increasing ability to offer credit has important sales and profit implications for marketers, just as the ability to obtain credit has important quality-of-life implications for consumers. However, despite the growth in credit availability, many con-

sumers are unable to gain access to the credit that they need and believe they deserve. The importance of this issue was recognized by Congress, which in 1974 passed the Equal Credit Opportunity Act prohibiting discrimination in the granting of credit on the basis of sex and marital status (ECOA 1975). In 1976 the Act was amended to include race, color, religion, national origin, receipt of income from a public assistance program, and age as proscribed characteristics. Further, in 1977, the Federal Trade Commission decided to devote a significant percentage of its then increased resources to the handling of all forms of credit abuse problems (*Advertising Age* 1977).

The federal legislation was directed largely at abuses in judgmental methods of granting credit. However, at that time judgmental methods that involve the exercise of individual judgment by a credit officer on a case-by-case basis were increasingly being replaced by a new methodology, credit scoring. William Fair has recently estimated that between 20 and 30% of all consumer credit decisions are now made by credit scoring, and that most of the very large credit granters including banks, finance com-

---

Noel Capon is Associate Professor of Business, Graduate School of Business, Columbia University. This paper is based in part on the author's testimony to the Subcommittee on Consumer Affairs, Committee on Banking, Housing and Urban Affairs, United States Senate.<sup>2</sup>

---

---

<sup>1</sup>Opening Statement of Senator Paul E. Tsongas (D., Mass.). See *Credit Card Redlining* 1979, p. 2.

<sup>2</sup>See *Credit Card Redlining* 1979, pp. 100-182.



panies, oil companies, retail merchants, and travel and entertainment cards now score their applicants (*Credit Card Redlining* 1979, p. 183-184).

This paper provides a critical analysis of credit scoring and may be viewed in part as a response to Nevin and Churchill's (1979) paper in this journal, which generally endorsed credit scoring systems. It will be shown that not only has their adoption led to major changes in the manner in which credit decisions are made but that these changes and the methodologies employed raise significant public policy issues.

## Credit Decision Methods

The conceptual framework for judgmental credit decisions has endured for many decades. This framework consists of the three "c's" of credit, character, capacity and capital, often joined by collateral and conditions, and indicated primarily by credit history and such other characteristics as income, occupation and residential stability. However, for such reasons as credit officer error, inconsistency in application of credit policies across credit officers, and high costs both in training and employing credit officers and in purchasing credit reports, innovative creditors have long sought more automated ways of making credit decisions.

Numerical scoring systems, first developed in the mail order industry in the 1930s and later used by large personal finance companies, were an attempt to address these concerns (Smalley and Sturdivant 1973, p. 229; Wonderlic 1952). In a typical system a number of predictor characteristics were chosen for their ability to discriminate between those who repaid their credit (goods) and those who did not (bads), and points were awarded to different levels of each characteristic. An individual applicant was judged on the relationship between his/her summated score across characteristics and independently set accept/reject cut-off values. Early systems employed such characteristics as occupation, length of employment, credit bureau clearance, personal references, marital status, bank account, neighborhood, collateral, length of residence, income, rent, life insurance ownership, sex and race. Although the Spiegel system (Smalley and Sturdivant 1973, p. 229) and a major study for the National Bureau of Economic Research (Durand 1941) used statistical procedures (one characteristic at a time) to determine the point assignments, most systems were based on trial and error.

Although the ability to make credit decisions on a quantitative rather than a judgmental basis represented an important advance, the widespread diffusion of quantitative methods did not occur until development of the necessary computer technology in the early 1960s. In computer-based systems, hereafter

termed *credit scoring systems*, the computational power of the computer is employed to identify, from a creditor's own historic files, those characteristics that best discriminate between the goods and the bads and to determine the point values for the various levels of each selected characteristic.

## Credit Scoring Systems: Development

The basic procedure for developing credit scoring systems involves the selection of samples of goods and bads from the creditor's files. Upwards of 50, and as many as 300 (Duffy 1977) potential predictor characteristics are obtained from the application blank. A multivariate statistical technique such as regression or discriminant analysis (see, for example, Beranek and Taylor 1976; Chatterjee and Barcun 1970; Long 1976; Myers and Forgy 1963) is employed, frequently in a stepwise manner, to identify those predictor characteristics, typically from eight to twelve, which contribute most to separation of the two groups. These selected characteristics, determined in part by the initial set of characteristics available from the application blank and in part by the data, and their point values are unique to an individual system. An example of a regionally based system of a major national retailer is shown in Table 1.

An applicant for credit is evaluated in a credit scoring system by simply summing the points received on the various application characteristics to arrive at a total score. This score may be treated in a number of ways depending on the system design. In the single cut-off method, the applicant's total score is compared to a single cut-off point score. If this score exceeds the cut-off, credit is granted; otherwise the applicant is rejected. More complex systems are based on a two-stage process. For example, the applicant's total score may be compared to two cut-off figures. If the score exceeds the higher cut-off, credit is awarded automatically, while if it falls below the lower cut-off, credit is automatically denied. If the score is between the two cut-offs, credit history information is obtained, scored, and the points added to the total score obtained from the application blank. If this new score is above a new higher cut-off, credit is awarded; if not, credit is denied.

The creditor sets his/her cut-off values on the basis of the probabilities of repayment and nonpayment associated with the various point scores and the trade-offs between type I and type II errors. The higher an acceptance cut-off is set, the lower the type I error (accepting applicants who fail to repay), while the lower a rejection cut-off value, the lower the type II error (failing to accept applicants who would have repaid).



**TABLE 1**  
**Major National Retailer's Final Scoring Table for Application Characteristics**

<u>Zip Code</u>		<u>Time at Present Address</u>	
Zip Codes A	60	Less than 6 months	39
Zip Codes B	48	6 months-1 year 5 months	30
Zip Codes C	41	1 year 6 months-3 years 5 months	27
Zip Codes D	37	3 years 6 months-7 years 5 months	30
Not answered	53	7 years 6 months-12 years 5 months	39
		12 years 6 months or longer	50
		Not answered	36
<u>Bank Reference</u>		<u>Time with Employer</u>	
Checking only	0	Less than 6 months	31
Savings only	0	6 months-5 years 5 months	24
Checking & Savings	15	5 years 6 months-8 years 5 months	26
Bank name or loan only	0	8 years 6 months-15 years 5 months	31
No bank reference	7	15 years 6 months or longer	39
Not answered	7	Homemakers	39
		Retired	31
		Unemployed	29
		Not answered	29
<u>Type of Housing</u>		<u>Finance Company Reference</u>	
Owns/buying	44	Yes	0
Rents	35	Other references only	25
All other	41	No	25
Not answered	39	Not answered	15
<u>Occupation</u>		<u>Other Department Store/Oil Card Major Credit Card</u>	
Clergy	46	Department store only	12
Creative	41	Oil card only	12
Driver	33	Major credit card only	17
Executive	62	Department store and oil card	17
Guard	46	Department store and credit card	31
Homemaker	50	Major credit card and oil card	31
Labor	33	All three	31
Manager	46	Other references only	0
Military enlisted	46	No credit	0
Military officer	62	Not answered	12
Office staff	46		
Outside	33		
Production	41		
Professional	62		
Retired	62		
Sales	46		
Semi-professional	50		
Service	41		
Student	46		
Teacher	41		
Unemployed	33		
All other	46		
Not answered	47		

Since the early 1960s the use of credit scoring systems has expanded enormously, as journals serving practitioners have been filled with articles extolling their virtues (e.g., Churchill, Nevin and Watson 1977a, b; Cremer 1972; Long and McConnell 1977; Main 1977; Myers 1962; Weingartner 1966). Further, passage of the Equal Credit Opportunity Act Amendments (*Federal Register* 1976) offered further endorsement of credit scoring systems when instructions regarding their use were specifically included in Reg-

ulation B, which implements the Act (*Federal Register* 1977).

In the hearings on the amendments creditors argued that adherence to the law would be improved if credit scoring systems were used. They contended that whereas credit decisions in judgmental systems were subject to arbitrary and capricious behavior by credit evaluators, decisions made with a credit scoring system were objective and free from such problems. Regulation B thus envisioned two categories of credit



decision systems, statistically sound and empirically derived credit scoring systems, and all others not satisfying the criteria of statistical soundness and empirical derivation, which are termed judgmental systems. This distinction has practical importance. For example, although age is a proscribed characteristic under the Act, if the system is statistically sound and empirically derived, it can be used as a predictive characteristic, provided that the elderly receive the maximum points awarded to any age category. The appropriate manner in which both types of systems should be used was spelled out in the Regulations.

Presently credit scoring systems are used extensively, especially among major credit granters. It is claimed that their use reduces bad debt losses, that more consumers are granted credit, and that organizational consistency in decision making is achieved. Further, the costs of granting credit are reduced, since less skilled personnel are required and fewer credit reports need be purchased (*Credit Card Redlining* 1979, p. 234-240; Fair, Isaac and Company 1977). However, despite the torrent of words endorsing credit scoring systems, when they are subject to detailed analysis many troubling issues of a consumer and public policy perspective can be identified.

### Analysis of Credit Scoring Systems: Variables and Points

The critical distinction between extant credit scoring systems and other methods of credit evaluation is the absence, in credit scoring, of an explanatory model. While judgmental systems are based, however imperfectly, upon a credit evaluator's explanatory model of credit performance, credit scoring systems are concerned solely with statistical predictability. Since prediction is the sole criterion for acceptability, any individual characteristic that can be scored, other than obviously illegal characteristics, has potential for inclusion in a credit scoring system. A partial list of characteristics used by creditors in the development of their systems is presented in Table 2. Few of these variables bear an explanatory relationship to credit performance. At best they might be statistical predictors whose relationship to payment performance can exist only through a complex chain of intervening variables. The overwhelming concern of creditors for prediction and a total unconcern for other issues was perhaps most tellingly demonstrated in the exchange between Senator Carl Levin (D., Michigan) and William Fair, chairman of Fair, Isaac and Company, the leading developer of credit scoring systems, at the Senate hearings on S15. Senator Levin asked Mr. Fair whether he *should be allowed* to use certain characteristics in the development of credit scoring systems (*Credit Card Redlining* 1979, p. 221):

**TABLE 2**  
**Partial List of Factors Used to Develop Credit Scoring Systems**

Telephone at home	Bank savings account
Own/rent living accommodations	Bank checking account
Age	Zip code of residence
Time at home address	Age of automobile
Industry in which employed	Make and model of automobile
Time with employer	Geographic area of U.S.
Time with previous employer	Finance company reference
Type of employment	Debt to income ratio
Number of dependents	Monthly rent/mortgage payment
Types of credit reference	Family size
Income	Telephone area code
Savings and loan references	Location of relatives
Trade union membership	Number of children
Age difference between man and wife	Number of other dependents
Telephone at work	Ownership of life insurance
Length of product being purchased	Width of product being purchased
First letter of last name	

Senator Levin: "You feel that you should be allowed to consider *race*?" (emphasis added)  
 Mr. Fair: "That is correct."  
 Senator Levin: "Would the same thing be true with *religion*?"  
 Mr. Fair: "Yes."  
 Senator Levin: "Would the same thing be true with *sex*?"  
 Mr. Fair: "Yes."  
 Senator Levin: "Would the same thing be true with *age*?"  
 Mr. Fair: "Yes."  
 Senator Levin: "The same thing be true with *marital status*?"  
 Mr. Fair: "Yes."  
 Senator Levin: "*Ethnic origin*?"  
 Mr. Fair: "Yes."

This exchange demonstrates very clearly that in the development of credit scoring systems, for Fair, Isaac and Company at least, no issue other than statistical predictability is of any consequence.<sup>3</sup> Although professing a commitment to obey the law, Fair, Isaac and Company, if statistical predictability were found and it were so able, would provide its customers with

<sup>3</sup>A logical extension of Mr. Fair's position would allow the inclusion of such characteristics as color of hair (if any), left or right-handedness, wear eyeglasses, height, weight, early morning drink preference (tea, coffee, milk, other), first digit of social security number, last digit of social security number, sexual preference (none, same, different, both), educational level, sports preference (football, baseball, tennis, soccer, golf, other), and favorite movie star (select from list), if it could be shown that they were statistically related to payment performance.



credit scoring systems that discriminated on the basis of *race, religion, sex, age, marital status* and *ethnic origin*.

The result, for consumers, of such a focus on prediction can be seen by examining two scoring tables which, in the author's experience, are typical of those in general use today. Table 1 presents the scoring table of the major national retailer. Of particular note are the following items:

- There are no economic variables such as income, debts, living expenses and the like.
- There are no variables for credit history.
- Zip code is a very important characteristic, and a "bad" residential location can put the applicant at a tremendous disadvantage.
- Applicants score fewer points if they rent their accommodations than if they own or are buying their home.
- The length of time the applicant has been at his/her present address or has been with his/her current employer are important characteristics. However, rather than greater residential and employment stability being worth an increasing number of points, as stability increases, the points awarded first decrease and then later increase.
- An applicant's occupation is an important characteristic. However, to be gainfully employed in the categories of driver, labor, or outside gains no more points than being unemployed.
- If the applicant fills out the application honestly and admits that he/she borrowed money from a finance company, he/she is severely penalized. Whether or not the loan was satisfactorily repaid is irrelevant.
- For many of the characteristics more points are awarded if the question goes unanswered than are awarded for many of the possible answers. Thus, the second most favorable way to score on the zip code characteristic is not to provide the information.

A second system, developed and used by the finance subsidiary of a major consumer durables manufacturer, is noted in Table 3. The following items are of interest:

- Income is an important variable. However, the points relationship does not increase monotonically; rather, the points fluctuate wildly as income increases.
- There are no variables for credit history.
- Applicants who own their own home score many more points than those with other arrangements.

- Increasing residential and employment stability are worth increasing numbers of points.
- The points awarded for age have a curvilinear relationship.
- Occupation is an important characteristic, and the unemployed category achieves the highest possible point score.
- Honestly providing a small loan reference results in being penalized.
- Many points are awarded for maintenance of either a checking or savings account, irrespective of the balances.

Though Congress embraced credit scoring systems, believing that their claimed objectivity offered advantages in enforcement of the Equal Credit Opportunity Act, the key goal of the Act was to:

“. . . establish(ed) as clear national policy that no credit applicant shall be denied the credit he or she needs and wants *on the basis of characteristics that have nothing to do with his or her creditworthiness* . . .” (*Equal Credit* 1976, p. 3) (emphasis added)

Congress insisted that creditors advise applicants of the reasons for adverse action since it was concerned with the educational value of such knowledge:

“. . . rejected applicants will now be able to learn where and how their credit status is deficient and this information should have a pervasive and valuable educational benefit . . .” (*Equal Credit* 1976, p. 4)

In identifying a set of proscribed characteristics (enumerated in ECOA), the clear intent of Congress was that acceptable characteristics are those that related to creditworthiness. While "relationship to creditworthiness" was not spelled out, many of the characteristics noted in Tables 1 and 3 do not evince a face valid relationship, for instance, those variables whose values are fluctuating—time at present address, time with employer (Table 1), and unpaid cash balance, age and income (Table 3). Given the concern for consumer education, it is difficult to believe that Congress would have accepted the fact that increased income (Table 3) and greater residential and employment stability (Table 1) should be regarded as indicators of *reduced* creditworthiness.<sup>4</sup>

Many other problems concerning the variables used and the points awarded exist with credit scoring systems. There is a real question of misleading the applicant. One might expect that provision of a financial reference would be reviewed positively, yet in both systems noted above, honesty is penalized. Also, there is the possibility that characteristics ac-

<sup>4</sup>The Senate Committee report asserts that: “. . . consumers particularly should benefit from knowing, for example, that the reason for their denial is their *short residence in the area, or their recent change of employment* . . .” (emphasis added).



tually employed act as surrogates for proscribed characteristics. Thus the Senate has heard testimony that a zip code acts as a surrogate for race (*Credit Card Redlining* 1979, p. 20–63, 261–264, 314–317). Accordingly, discrimination can result when zip code is used as a predictor characteristic, when different cut-off values are employed for different zip codes, or when credit scoring systems are developed at the individual zip code level. Differential treatment of types of income, such as that from part-time employment, alimony, child support and separate maintenance payments, discriminates against women. Furthermore, own/rent accommodation may discriminate against minorities as a result of historical discrimination in granting of mortgage loans, just as occupation and length of time with employer may discriminate against women because of historic employment practices and reduced employment stability due to pregnancy and childbearing, respectively. In the same way, age of automobile may discriminate against the handicapped.<sup>5</sup>

Since credit history information only enters credit scoring systems at a second stage, if at all, many applicants are denied credit despite the fact that they had excellent credit records (Chandler and Ewert 1975; *Credit Card Redlining* 1979, p. 63–70). Their reputations are unjustly injured, and severe psychological trauma may also ensue (*Credit Card Redlining* 1979, p. 135–136). The use of mere statistical prediction to make decisions may violate the constitutional guarantees of the equal protection and due process clauses of the 5th and 14th Amendments (*Credit Card Redlining* 1979, p. 137–138). The equal protection clause addresses the question of making decisions on individuals on the basis of characteristics that are both “irrelevant and unchangeable,” while due process states that “individual cases must be decided on their own merits.” In passing ECOA, Congress proscribed characteristics that were either immutable (race, color, national origin, sex) or central to the individual’s life (religion, marital status). Characteristics still frequently employed in credit scoring systems such as number of dependents, age, occupation and place of residence appear to have many similarities to these proscribed characteristics, both in terms of being “irrelevant and unchangeable” and having little or nothing to do with “merit” in the case of a credit decision.<sup>6</sup>

<sup>5</sup>Whether or not such surrogate variables could legally be employed in a credit scoring system would depend upon the results of application of an “Effects Test.” See *Griggs v. Duke Paper Co.*, 401 U.S. 424 (1971), and *Albermarle Paper Co. v. Moody*, 422 U.S. 405 (1975).

<sup>6</sup>Nevin and Churchill (1979) present empirical support for the proposition that if characteristics correlated with proscribed characteristics were disallowed in credit scoring systems, the predictive ability of the model would be reduced. On the basis of an example in which “we

## Analysis of Credit Scoring Systems: Development

The focus of the previous section was on problems involving the selection of predictor characteristics and the award of point values. In this section a series of methodological issues in the development of credit scoring systems is addressed. It will be shown that there are real questions as to whether credit scoring systems satisfy the legal requirements of empirical derivation and statistical soundness. The areas of concern are several and are discussed below.

### Bias

The correct way to develop a credit scoring system is to sample randomly an historic applicant population. Creditors typically do not sample in this manner, however, for only data from those applicants previously awarded credit can provide samples of goods and bads. Since a considerable percentage of applicants was historically denied credit, systems based only on a population of accepted applicants where there is a corresponding population of denied applicants must be biased. Indeed, it has been shown that not only are biased estimates obtained, it is not possible to estimate in which direction the bias lies (Avery 1977). This problem is more severe in those systems that were originally developed before enactment of ECOA, when variables that are now illegal were used to make credit decisions. Despite revalidation, these systems are both biased and contaminated by illegal discrimination.

Developers of credit scoring systems are aware of the problem of using biased samples and have developed techniques in attempts to solve it. In the augmentation method, a sample of denied applicants is separated into goods and bads on the basis of the relationship of their application characteristics to those of the actual goods and bads. The actual and denied goods are then grouped, as are the actual and denied bads, and the credit scoring system is developed from the augmented sample. However, as Shinkel (1977) has shown, biased estimates are still obtained with this and alternative procedures.<sup>7</sup>

have tried to make the assumptions realistic with respect to industry experience” (p. 102), they show that not only would the fictional consumer finance company earn less profit, fewer applicants would be awarded credit, using a restricted model. They fail to note that a profit maximizing finance company should award credit to all applicants, in which case profits would be \$1.9M versus \$1.78M and 9,000 versus 6,019 “good” applicants would be granted credit.

<sup>7</sup>Eisenbeis (1978) has discussed a number of statistical problems relating to the use of discriminant analysis in credit scoring. They include violations of the assumption about the underlying distributions of the characteristics, use of linear instead of quadratic discriminant functions when group dispersions are unequal, difficulty in demonstrating the significance of each characteristic included in the system, and estimation of classification error rates.



### Multicollinearity

Credit scoring systems are developed from a large group of contender characteristics. In stepwise procedures the characteristic that explains the greatest variance enters the discriminatory function first, followed by other characteristics which in turn explain the greatest residual variance. However, there is no requirement that despite their ability to explain residual variance, subsequently entered variables are not correlated with variables previously entered. Thus, the coefficients of variables entered early to the equation are continually modified as successive variables are entered. The final point values assigned are far from being a true reflection of the discriminatory power of the single variable and are contaminated by a host of intercorrelations (Hsia 1978). A variable with good predictive ability but highly correlated to an entered variable will not enter the final equation. No greater concern for multicollinearity is shown in systems where the characteristics are preselected.<sup>8</sup>

An associated problem of intercorrelation of variables arises in the development of the second stage of two-stage systems in which the potentially discriminating credit history variables act only on the residual variance. Because of the intercorrelation between credit history variables and those variables already entered, the effect of credit history is severely circumscribed.

### Sample Size

Credit scoring systems are frequently developed with insufficiently large samples to achieve reliability in the assignment of point values. Thus, for the occupation characteristic of a credit scoring system employed by a major oil company, the occupations of farm foremen and laborers, enlisted personnel, clergymen, entertainers, farmers and ranchers, and government and public officials received few points. However, the sample sizes on which the point scores are based were, respectively, three, twenty-three, four, four, three and three. The point values are clearly unreliable. Similar patterns occur when zip code is used as a characteristic. Thus, for a regional trading area with hundreds of zip codes, the use of sample sizes of 3,000 or fewer subjects results in the point scores for many zip codes being based on very few data points. The system described in Table 3 was developed from a mere 640 data points (which may in part explain the strange income relationship).<sup>9</sup>

<sup>8</sup>The multicollinearity problem could perhaps be addressed by factor analysis and the use of factor scores. However, such a procedure would run into the problem of a legal requirement to disclose reasons for adverse action, where the "reason" would now be a factor score correlated to a greater or lesser extent with many original variables.

<sup>9</sup>The zip code analysis for the Table 1 system was based on between 500 and 600 individual zip codes, which, at an estimated maximum

### Judgmental Aggregation

The empirical requirement for credit scoring systems is violated when credit scorers attempt to overcome the reliability problem. Then they aggregate individual units of a variable but in a nonempirical, arbitrary manner. The geographic unit, for example, may be defined not as a small unit such as zip code but as a state or regional grouping of states under no rationale other than, perhaps, geographic contiguity.<sup>10</sup> In the system described in Table 1, the 20 gross occupation categories were developed from 300 or more fine-level occupations (*Credit Card Redlining* 1979, p. 166-168).

Not only are the occupation categories developed in an arbitrary manner, they are not a mutually exclusive set: an individual applicant could be assigned to a number of different categories. Thus, for example, a sales manager could be assigned as executive (62 points), manager (46 points), office staff (46 points), professional (62 points) or sales (46 points).

**TABLE 3**  
Final Scoring Table for Finance Subsidiary of Consumer Durables Manufacturer

<b>Unpaid Cash Balance</b>		<b>Age</b>	
\$ 0-299	26	26-29	5
300-499	16	30-34	0
500-599	20	35-39	4
600-699	15	40-49	9
700 & above	4	50-54	14
		55 & above	17
<b>Time at Present Address</b>		<b>Income (monthly)</b>	
Less than 1 year	4	\$ 0-599	37
1-2 years	6	600-699	47
3-9 years	8	700-799	40
10 years or longer	10	800-899	36
		900-1,099	44
		1,100-1,299	39
		1,300 and above	49
<b>Time with Present Employer</b>		<b>CoApplicant</b>	
Less than 1 year	2	Employed	6
1-2 years	10		
3-5 years	12		
6-9 years	16		
10 years or longer	22		
<b>Residence</b>		<b>Financial</b>	
Own	17	Major credit card	22
Rent or live with relative	0	Small loan reference	(7)
		No checking or savings account	(18)

sample size in the 3,000 to 5,000 range, implies that many zip codes contained very few data points.

<sup>10</sup>For a worked example of the problems of aggregation with geographic units, see *Credit Card Redlining*, p. 122-125. Also see p. 384-86 for a discussion of aggregation and homogeneity problems in the use of zip codes.

TABLE 3 (continued)

**Occupation<sup>a</sup>****Group 1-19**

Accountants, Auditors	Machinists
Architects, Designers	Physicians, Dentists
Bank tellers/clerks	Pilots (nonmilitary)
Business executives	Postal employees
College professors	Real estate personnel
Computer programmers	Reporters, newsmen
Engineers, Chemists	Salesmen (not
Factory inspectors	department store)
Factory workers	Supervisors, nonoffice
(semi-skilled)	Supervisors, office
Farm owners	Systems analysts
Field representatives	Teachers, instructors
Firemen, Rangers	Unemployed
Insurance agents,	
Appraisers	

**Group 2-13**

Building superintendents	Office managers
Carpenters, Craftsmen	Plumbers, Pipefitters
Clergymen	Policemen, Detectives
Clerical workers,	President/Owner of small
Bookkeepers	firm
Computer operators	Printers, Pressmen
Electricians	Railroad employees
Foremen, factory	Registered nurses
Government employees	Repairmen
Guards	Sales clerks
Installers	Seamen (nonmilitary)
Lawyers, Judges	Secretaries,
Maintenance men	Stenographers
Managers, other than	Shipping and stock clerks
office	Stewards, Stewardesses
Mechanics	Taxi drivers, Chauffeurs
Medical and Dental	Technicians, Researchers
assistants	

**Group 3-0**

Artists, Entertainers,	Hotel and Restaurant
Writers	employees
Assemblers	Laborers (unskilled)
Bartenders	Machine operators
Construction workers	Painters
Contractors, Builders	Social workers
Cooks, Bakers, Butchers	Tailor/Seamstress
Delivery and Route men	Truck and Bus drivers
Dishwashers	Waiter/Waitress
Domestics, Janitors	Warehousemen
Heavy equipment	Welders
operators	
Hospital employees	
(unskilled)	

**Group 4-0**

Other—Not directly related to Groups 1, 2 or 3 above

<sup>a</sup>All self-employed should be investigated

A U.S. Senator might be classified as executive (62 points), professional (62 points), manager (46 points) or all other (46 points).

**Judgmental System Constraints**

Since the methodology used to develop credit scoring systems is brute force empiricism, point value assignments to levels of characteristics in the final scoring table are often absurd, as indicated in the previous section. To overcome the consequent problems of credit scoring personnel ignoring the system, developers impose constraints on point assignments a priori (Churchill, Nevin and Watson 1977b; Fair, Isaac and Company 1977). While final scoring tables may thus be less absurd than otherwise, the impact of this procedure is to violate the empirical requirement of ECOA.

**Overriding**

The overriding procedure is also a violation of the empirical requirement. Overriding exists when a declined applicant calls to complain and, either on the basis of no information other than the protest or on the basis of some extra information, the decision is reversed and credit is awarded. Not only is use of the overriding procedure a statement that the system is not doing the job it was designed to do, it is discriminatory procedure against those who are less vociferous following credit denial.

**Histogram Error**

When continuous characteristics such as time are used, serious errors may be introduced to the scoring table by using a series of discrete categories rather than the underlying continuous characteristic. Thus, for the characteristic "time at present address" in the scoring system described in Table 1, there are a series of histogram errors. For instance, an applicant with a residency of seven years and five months scores 30 points and one month later scores 39 points, a "present" of nine points. Conversely, a person with a residency of five months scores 39 points and one month later "loses" nine points. Errors of over 25% misclassification have been noted because of this histogram effect (The Sorites Group 1978).

In this section, seven areas of methodological concern have been noted. Not only were troubling statistical issues raised, it was shown that the procedures employed for the development of credit scoring systems may violate the legal requirements of ECOA that they be empirically derived and statistically sound. Certainly, the institution of careful procedures may obviate some problems, for example, overriding, but fatal methodological flaws may render some insoluble.



## Discussion

In this paper the background, development and rationale of credit scoring systems have been described. The benefits to creditors of such systems have been so vigorously promoted that regulations concerning their use have been specifically written into the law, and their adoption has been extremely widespread, especially among major creditors. Thus, within the past 20 years a major change in credit granting practice affecting millions of consumers has occurred in the United States. However, public debate has been virtually absent on this topic.

This paper redresses the balance and focuses a critical eye on credit scoring systems. When subject to intensive examination, a very different picture emerges from that portrayed by the multitude of credit scoring boosters.

An examination of the development of credit scoring systems reveals a host of statistical issues that may pose severe legal problems for creditors. Statisticians have only recently begun to investigate these systems, yet their early findings are very troubling. It is perhaps not unlikely that 20 years of intensive study of these systems paralleling the 20 years of development just past may lead to conclusions even more serious than are justified by our present knowledge.

The more troubling aspect, however, has less to do with statistical issues than with conceptual ones. The brute force empiricism that characterizes the development of credit scoring systems leads to a treatment of the individual applicant in a manner that offends against the traditions of our society. When predictive decisions regarding individuals have to be made, they are based typically on variables that bear an explanatory rather than a statistical relation to the behavior being predicted, notably the actual historic performance in a similar or related area. For instance, job promotion rests heavily on job performance; selection for college is based on high school grades and aptitude tests. Yet credit scoring developers use any characteristic that discriminates as long as they can get away with it; they have even used the first letter of a person's last name.<sup>11</sup>

---

<sup>11</sup>Disclosure by Morton Schwartz, General Credit Manager, J. C. Penney Company, at a meeting of the Trustees of the Credit Research Center, Atlanta, Ga., November 10, 1977.

As far as individuals not yet in the credit marketplace are concerned, who have no credit history, the characteristic of extant systems that they ignore credit history is no argument for their use. It is arbitrary and unfair to make decisions on these applicants on the basis of points awarded arbitrarily for the characteristics of those already in the market. Experience of enterprising retailers suggests that a system characterized by low initial credit limits and tight controls is a better way to treat new applicants.

What is needed, clearly, is a redirection of credit scoring research efforts toward development of explanatory models of credit performance and the isolation of variables bearing an explanatory relationship to credit performance. Such variables are likely related to economic factors (ability to pay) and credit history factors (demonstrated willingness to pay). In present systems, economic factors do not always enter the credit scoring tables, in part because they are highly correlated with other entering variables, for instance, zip code and income. Furthermore, since creditors are unwilling to pay the cost of credit reports, credit history factors are relegated to the second stage and their use is thus minimized, despite ample evidence that they provide the strongest relationship to future credit performance (Chandler and Ewert 1975; *Credit Card Redlining* 1979, p. 376; Long and McConnell 1977).

It is, of course, possible that well-developed explanatory models would be less predictive overall and more costly to implement than currently employed credit scoring systems. Even if this were true, and it may not be so, such increased costs should be weighed against the social cost of employing systems such as those described in this paper that provide a dispassionate observer with "a chilling experience."

This analysis should not be construed as advocacy for traditional judgmental systems nor as argument against the thrust toward objectivity and consistency in credit decision making. Such a direction is clearly a positive one. What is critical, however, is treatment of the individual in a fair and just manner and his/her protection from arbitrary treatment. The individual should be judged on characteristics that are ultimately related to the decision under consideration; brute force empiricism has no place in decisions of such importance to individual citizens.

---

## REFERENCES

*Advertising Age* (1977), "FTC's Newest Budget Boosts Credit Probes: Regional Ad Watch Cut" (January 3), 2.  
Avery, Robert B. (1977), "Credit Scoring Models with Dis-

criminant Analysis and Truncated Samples," unpublished paper, Graduate School of Industrial Administration, Carnegie-Mellon University.



- Beranek, William and Walton Taylor (1976), "Credit Scoring Models and the Cut-Off Point—A Simplification," *Decision Sciences*, 7 (July), 394-404.
- Board of Governors of the Federal Reserve System (1976a), *Banking and Monetary Statistics, 1941-1970*, Washington, DC: U.S. Government Printing Office.
- (1976b), *Annual Statistical Digest, 1971-1975*, Washington, DC: U.S. Government Printing Office.
- (1980), *Annual Statistical Digest, 1974-1978*, Washington, DC: U.S. Government Printing Office.
- Chandler, Gary G. and David C. Ewert (1975), "Cost and Value of Credit Bureau Reports," *Proceedings of the Eastern Finance Association*, 71-72.
- Chatterjee, Samprit and Seymour Barcun (1970), "A Non-parametric Approach to Credit Screening," *Journal of the American Statistical Association*, 65 (March), 150-154.
- Churchill, Gilbert A. Jr., John R. Nevin and R. Richard Watson (1977a), "The Role of Credit Scoring in the Loan Decision," *The Credit World* (March), 6-10.
- , ——— and ——— (1977b), "Developing a Credit Scoring System," *The Credit World* (April), 6-10.
- Credit Card Redlining* (1979), Hearings before the Subcommittee on Consumer Affairs of the Committee on Banking, Housing and Urban Affairs, United States Senate, 96th Congress, First Session, on S15, June 4 and 5, 1979, Washington, DC: U.S. Government Printing Office.
- Cremer, Richard E. (1972), "Consumer Credit Program Key to Montgomery Ward's Growth," *Direct Marketing* (September), 40-50.
- Duffy, William (1977), "The Credit Scoring Movement," *Credit* (September), 28-30.
- Durand, David (1941), *Risk Elements in Consumer Installment Financing*, Financial Research Program, Studies in Consumer Installment Financing, 8, New York: National Bureau of Economic Research.
- Eisenbeis, Robert A. (1978), "Problems in Applying Discriminant Analysis in Credit Scoring Models," *Journal of Banking and Finance*, 2 (October), 205-219.
- Equal Credit Opportunity Act* (1975), U.S.C., Title 15, Sec. 1691 et seq.
- Equal Credit Opportunity Act Amendments of 1976* (1976), Report of the Committee on Banking, Housing and Urban Affairs, United States Senate, 94th Congress, Second Session, Washington, DC: U.S. Government Printing Office.
- Fair, Isaac and Company (1977), "Notes to Accompany Credit Scoring Seminar," unpublished manuscript, San Rafael, California.
- Federal Register* (1976), "Part 202—Equal Credit Opportunity," 41, No. 216 (Monday, November 8), 49131-49.
- (1977), "Part 202—Equal Credit Opportunity," 42, No. 4 (Thursday, January 6), 1242-1263.
- Hsia, David C. (1978), "Credit Scoring and the Equal Credit Opportunity Act," *The Hastings Law Journal*, 30 (November), 371-448.
- Long, Michael S. (1976), "Credit Screening System Selection," *Journal of Financial and Quantitative Analysis*, 11 (June), 313-328.
- and John J. McConnell (1977), "Credit Scoring System Aids Second Mortgage Makers," *Mortgage Banker* (July), 20-36.
- Main, Jeremy (1977), "A New Way to Score with Lenders," *Money* (February), 73-74.
- Myers, James H. (1962), "Numerical Scoring Systems for Retail Credit Evaluation," *The Credit World*, 50 (April), 5-7.
- and Edward W. Forgy (1963), "The Development of Numerical Credit Evaluation Systems," *Journal of the American Statistical Association*, 58 (September), 799-806.
- Nevin, John R. and Gilbert A. Churchill, Jr. (1979), "The Equal Credit Opportunity Act: An Evaluation," *Journal of Marketing*, 43 (Spring), 95-104.
- Shinkel, Bernard (1977), "Alternative Methodologies to Reduce Selection Bias in Credit Scoring," paper presented at the Financial Management Association Meeting, Seattle, Washington (October 14).
- Smalley, Orange A. and Frederick D. Sturdivant (1973), *The Credit Merchants: A History of Spiegel, Inc.*, Carbondale, IL: Southern Illinois University Press.
- The Sorites Group (1978), "Credit Scoring and the Law," unpublished manuscript, The Sorites Group, Inc., Springfield, Va.
- Weingartner, H. Martin (1966), "Concepts and Utilization of Credit Scoring Techniques," *Banking*, 58 (February), 51-53.
- Wonderlic, E. F. (1952), "An Analysis of Factors in Granting Credit," *Indiana University Bulletin*, 50 (no. 27), 163-176.