

South Dakota State University

## Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange

---

Electronic Theses and Dissertations

---

1974

### Income Expectations Among ADC Recipients in Rapid City, South Dakota : An Application of Discriminant Analysis

Craig Edwards Clark

Follow this and additional works at: <https://openprairie.sdstate.edu/etd>

---

#### Recommended Citation

Clark, Craig Edwards, "Income Expectations Among ADC Recipients in Rapid City, South Dakota : An Application of Discriminant Analysis" (1974). *Electronic Theses and Dissertations*. 4677.  
<https://openprairie.sdstate.edu/etd/4677>

This Thesis - Open Access is brought to you for free and open access by Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. For more information, please contact [michael.biondo@sdstate.edu](mailto:michael.biondo@sdstate.edu).

INCOME EXPECTATIONS AMONG ADC RECIPIENTS  
IN RAPID CITY, SOUTH DAKOTA-AN  
APPLICATION OF DISCRIMINANT  
ANALYSIS

BY

CRAIG EDWARDS CLARK

This thesis is approved as a reasonable and independent investigation by the Department of Economics, South Dakota State University, and is acceptable in meeting the requirements for this degree. Acceptance of this thesis does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

*[Signature]*  
Thesis Advisor Date

A thesis submitted  
in partial fulfillment of the requirements for the  
degree Master of Science, Major in  
Economics, South Dakota  
State University

1974



## ACKNOWLEDGMENT

The author wishes to express sincerest gratitude to Dr. Charles Vanderziel for his invaluable assistance in the design and development of the present study. Special thanks is also due to H. Stanley Bolding for his efforts in providing assistance in the operation of the computer programs used to analyze the study's data. The author is most grateful for the efforts of Ms. Carole Anderson in the editing and typing of the manuscript.

Finally, this study could never have reached completion without the assistance of the Department of Social Services, Pennington County, South Dakota, and the fifty-one families surveyed in this study who gave generously of their own time in the hope of bringing to light the common affliction shared by all: poverty.

|   |    |
|---|----|
| Determining Statistical Distributions   | 14 |
| Evaluating the Variability              | 17 |
| Classifying Individuals into Groups     | 17 |
| THE WELFARE POPULATIONS                 | 19 |
| Defining the AOC Population             | 19 |
| THE CHOICE OF VARIABLES                 | 23 |
| ANALYSIS OF THE RESULTS                 | 28 |
| RELEVANCE OF VARIABLES TO THE CHOICE OF | 28 |
| STATISTICAL SIGNIFICANCE OF THE RESULTS | 31 |

## TABLES OF CONTENTS

| Chapter | TABLES OF CONTENTS                                    | Page |
|---------|---|------|
| 3.      | CLASSIFICATION OF INDIVIDUALS BY GROUP.               | 32   |
|         | LIST OF TABLES . . . . .                              | vi   |
|         | LIST OF FIGURES. . . . .                              | vii  |
| <br>    |   |      |
| Chapter |   |      |
| 1.      | INTRODUCTION. . . . .                                 | 1    |
|         | INCOME EXPECTATIONS AND POVERTY . . . . .             | 3    |
|         | THE SPECIFIC PROBLEM. . . . .                         | 5    |
|         | THE RESEARCH OBJECTIVES . . . . .                     | 6    |
|         | CONTENT OF SUCCEEDING CHAPTERS. . . . .               | 6    |
| 2.      | FORMULATION OF THE MODEL. . . . .                     | 8    |
|         | THE ANALYTIC TECHNIQUE. . . . .                       | 10   |
|         | Assumptions of the Model . . . . .                    | 10   |
|         | The Linear Classification Procedure. . . . .          | 12   |
|         | APPLICATIONS OF THE DISCRIMINANT<br>FUNCTION. . . . . | 14   |
|         | Determining Statistical Distance . . . . .            | 14   |
|         | Evaluating the Variables . . . . .                    | 17   |
|         | Classifying Individuals into Groups. . . . .          | 17   |
|         | THE WELFARE POPULATIONS . . . . .                     | 18   |
|         | Delimiting the ADC Population. . . . .                | 19   |
|         | THE CHOICE OF VARIABLES . . . . .                     | 23   |
| 3.      | ANALYSIS OF THE RESULTS . . . . .                     | 28   |
|         | RELEVANCE OF VARIABLES TO THE MODEL . . . . .         | 28   |
|         | STATISTICAL SIGNIFICANCE OF THE MODEL . . . . .       | 31   |

| Chapter   | Page |
|---|------|
| 3. CLASSIFICATION OF INDIVIDUALS BY GROUP. . . . .            | 32   |
| EVALUATING THE VARIABLES. . . . .                             | 35   |
| Evaluating the Income Variable . . . . .                      | 37   |
| Evaluating the Non-Income Variables. . . . .                  | 40   |
| The Desire to Leave ADC. . . . .                              | 45   |
| 4. SUMMARY AND IMPLICATIONS, AND<br>RECOMMENDATIONS . . . . . | 50   |
| IMPLICATIONS OF THE STUDY . . . . .                           | 50   |
| RECOMMENDATIONS FOR FURTHER RESEARCH. . . . .                 | 52   |
| BIBLIOGRAPHY . . . . .  | 54   |
| APPENDIX . . . . .  | 58   |

## LIST OF TABLES

| Table |  | Page |
|-------|--|------|
| 1.    | Matrix of Correlation Between<br>Discriminant Variables. . . . .                                   | 30   |
| 2.    | Misclassification Probabilities . . . . .  | 35   |
| 3.    | Ranking of Discriminant Coefficients. . . . .  | 38   |
| 4.    | Percent Contribution of Adjusted<br>Discriminant Variables to Total<br>Discriminant Score. . . . . | 42   |
| 5.    | Value of Group Means, Discriminant<br>Coefficients, and $D^2$ Statistic. . . . .                   | 46   |
| 6.    | Pooled Dispersion Matrix<br>(Symmetric Matrix). . . . .  | 62   |

## LIST OF FIGURES

| Figure |                      | Page |
|--------|----------------------|------|
| 1.     | Figure 3-1 . . . . . | 33   |
| 2.     | Figure 3-2 . . . . . | 33   |

erations related to low birth rates, the aging population, and the depletion of social resources, especially those of mental and physical health, and so forth. To the extent that these factors are likely to bring to mind the name of the young economist, or Edgeworth.

Welfare economics, as a branch of economic theory, represents a rather recent application of the economic theory to certain contemporary social problems, especially the problem of poverty. Hence, welfare economics is not strictly social welfare economics, but has, in fact, been a part of economic theory but a system of welfare economics that provides economic support for the aged, the disabled, or handicapped or elderly. The cost of such a system, stagger- ing cost of 20 billion dollars annually.

So it is welfare economics, especially, that has with the advent of the great industrial and technological societies began to receive an important and perhaps always self-sustaining. The first edition of the book of John Maynard

---

<sup>1</sup>1975 Statistical Abstract of the United States; Table 467, p. 228.



## Chapter 1

### INTRODUCTION

Welfare economics is one of the oldest areas within the body of economic thought. It comprises a spectrum of considerations related to individual utility functions, optimization of social resources, marginal rates of social transformation, and so forth. To the economist, welfare economics would likely bring to mind the names of Antonelli, Walras, Fisher, or Edgeworth.

Welfare economics, to the man on the street, represents a rather recent application of modern economic theory to certain contemporary social problems, specifically the problem of poverty. Hence, welfare economics, or more accurately social welfare economics, not only represents a body of economic theory but a system of welfare programs and payments that provide economic support for over 40 million indigent, disabled, or handicapped or elderly American citizens at a rather staggering cost of 20 billion dollars annually.<sup>1</sup>

Social welfare economics began in earnest with the advent of the great depression. Free industrialized societies began to realize unregulated economies were not always self-sustaining. The depression and the work of John Maynard

---

<sup>1</sup>1973 Statistical Abstract of the United States; Table 462, p. 288.

Keynes, which provided the theoretical link between consumption, income, and aggregate demand, created a favorable climate for the growth of a social welfare system. Individuals in favor of welfare spending saw it as a stabilizing device in which payments would increase or decrease with business cycles to stimulate or retard growth in aggregate demand. Welfare spending would be a convenient faucet to turn up or down as dictated by economic (and sometimes political) necessity. Since that time, most Americans have witnessed the growth of a variety of social welfare programs such as Social Security, Medicare, Medicaid, and Aid to Dependent Children, and aid to the blind and disabled, to mention a few.

With the growth of a permanent social welfare system, it is not surprising that a multitude of books, papers, and studies on the socio-economic problems of poverty, e.g., low levels of education and training, high rates of unemployment, prejudice, geographic isolation, and an apparent lack of motivation or desire for economic self-improvement have been published. Of all the aforementioned problems, the lack of motivation and the absence of positive economic expectations have posed serious questions for the social welfare system, especially in light of growing evidence which suggests the existence of a permanent class of welfare recipients where literally generations of the same family become dependent upon state and federal assistance.

The question of how economic expectations and motivations of an individual are related to his status as a welfare recipient is important to the existence of many current welfare programs, especially if these programs contribute significantly to the perpetuation of individuals in the social welfare system. A more fundamental question is what specific socio-economic factors contribute most to the economic expectations of the individual recipient and how do these factors affect the specific economic behavior of the welfare recipient.

In the present study the question of expectations of welfare recipients, specifically income expectations, is addressed. The general population chosen for examination are women and children receiving Aid to Dependent Children (ADC) payments. This general population is examined in an attempt to determine the specific factors contributing to positive or negative income expectations. These factors cover a wide range of socio-economic considerations, such as employment, receivable income, race, age of recipients, the number of dependent children, and many other factors.

#### INCOME EXPECTATIONS AND POVERTY

There are several reasons for examining income expectations of the poor. For one, income expectations should offer some evidence of the availability of employment for welfare mothers. This follows because receivable monthly welfare income is more or less constant. If employment serves as a

major avenue toward higher income, income expectations may represent a measure of meaningful employment opportunities. Similarly, income expectations may represent a measure of the work incentive present in a particular state's welfare program. Some states offer rather significant incentives to work in the form of additional work related income. Thus, since income expectations relate to employment and the desire for employment may be related to the existence of employment incentives in the welfare program, income expectations may represent the ability of a state welfare system to encourage economic self-sufficiency.

Income expectations may also relate to the economic behavior of the poor. Families with differing income expectations will likely differ in consumer behavior. A family with strong expectations of higher income may be expected to demand a wider range of goods and services. This same family may set aside savings with the expectation of increasing future consumption. It is even possible that the expectation of higher income will encourage the struggle for leaving ADC altogether.

A family with negative income expectations may be expected to pursue a different course. Given the guarantee of a steady monthly welfare income, the family with negative income expectations may completely ignore any existing economic alternatives that could ultimately help their situation.

Income expectations may offer evidence to support the assertion that certain cultural or racial factors bear heavily on the probability of a welfare family to control its own economic destiny. If one particular ethnic or cultural group exhibits a strong tendency toward low or high income expectations, it is likely that the particular racial or cultural characteristics of the group contribute heavily to income expectations and other expectations as well. It would not be surprising to discover the dominance of income expectations in determining the bulk of a welfare recipient's economic self-perception, both present and future.

Thus, income expectations may represent the culmination of a number of factors influencing the outlook, behavior, and motivation of the poor. Certainly, a closer look at the influences underlying the income expectations of the poor is in order.

#### THE SPECIFIC PROBLEM

The specific problem may be stated as follows: Does an identifiable, identical set of variables exist among unmarried ADC recipients in Rapid City, South Dakota, which can serve to differentiate between recipients with positive and negative income expectations?

The specific problem statement may be clarified by stating that the intent of the research is to identify a

common set of influences shared by each recipient which either contribute to positive or negative income expectations.

### THE RESEARCH OBJECTIVES

The present research effort has three major objectives. They are:

- (1) To estimate whether or not the hypothesized set of underlying variables serves as a basis for discriminating between groups with differing income expectations.
- (2) To test the statistical significance of differences between groups with differing income expectations, based on a linear combination of the hypothesized set of variables.
- (3) To assess the relative contribution of each variable to separating groups with differing income expectations.

### CONTENT OF SUCCEEDING CHAPTERS

The historical and theoretical foundations for discriminant analysis, the analytic technique used in the study, are presented in Chapter 2. Because this study attempts to discriminate between welfare recipients with positive or negative income expectations, Chapter 2 includes a discussion of how discriminant analysis provides a specific measure of group differences. An effort is made to suggest several possible applications of certain underlying theoretical propositions to the conclusions stated in Chapter 3.

The question of the statistical results of the study are addressed in Chapter 3. Several techniques are presented to assess the ability of the model to discriminate between sample groups. In addition, the socio-economic characteristics of the sample groups are discussed with regard to their influence on income expectations. The statement of objectives contained in the present chapter are confronted in Chapter 4. Avenues for further research are discussed and the normative aspects of research in welfare economics are explored. Finally, the universal applicability of such studies are discussed.

The identification of the category or group to which an individual or object belongs on the basis of its observed characteristics is referred to as Classification. When the observed characteristics are numerical measurements, the procedure is referred to as Discrimination.<sup>1</sup> Hence, the thrust of this research is the identification and specification of a discriminant function that will provide for the separation of sample groups.

The use of discriminant functions is by no means a new or novel idea. This technique has been used for nearly a half century in the fields of biology, psychology, education, and medical research. Since the technique allows for the separation or classification of individuals into groups, it is

<sup>1</sup>David L. Sells, *International Encyclopedia of Social Sciences* (New York: The MacMillan Co. Free Press, 1968), p. 533.

## FORMULATION OF THE MODEL

In order to confront the objectives specified in Chapter 1, a method must be identified to assist in distinguishing sample groups on the basis of certain shared characteristics. This separation of sample groups should be identified in a specific functional relationship. Ideally, a function of this type will provide a foundation for predicting and evaluating group membership based upon the influence of the set of shared characteristics.

The identification of the category or group to which an individual or object belongs on the basis of its observed characteristics is referred to as Classification. When the observed characteristics are numerical measurements, the procedure is referred to as Discrimination.<sup>1</sup> Hence, the thrust of this research design is the identification and specification of a discriminant function that will provide for the separation of sample groups.

The use of discriminant functions is by no means a new or novel idea. This technique has been used for nearly a half century in the fields of biology, psychology, education, and medical research. Since the technique allows for the separation or classification of individuals into groups, it is

---

<sup>1</sup>David L. Sills, International Encyclopedia of Social Sciences (New York: The McMillan Co. Free Press, 1968), p. 553.



not difficult to appreciate the growing interest in this technique.

Discriminant analysis has recently been applied to such mundane considerations as classifying individuals into groups of consumers likely to buy a certain brand of automobile or to apply for a certain credit card. The technique has been used in the field of education to discriminate between students likely to achieve academic success. However, it has been within the sciences that the earliest applications of discriminant analysis were understood. It was the work of the eminent biologist R. A. Fisher in the classification of plant species on the basis of petal characteristics that was to be one of the earliest and most profound applications of discriminant analysis.

Until the 1930's, the use of discriminant analysis was confounded with the problem of testing the equality of two or more distributions and determining a test statistic designed to test the equality of these distributions.<sup>2</sup> Fisher's work, however, directly confronted the problem of discrimination: that of identifying a function or rule which classifies an observation into a specific category or group through a linear combination of component variables. Fisher's work established the analytic technique necessary for optimal separation of groups. He thereby provided a means for classification and

---

<sup>2</sup>Somesh Das Gupta, "Theories and Methods in Classification: A Review," in Discriminant Analysis Applications, ed. by T. Cacoullas. (New York: Academic Press, 1973) p. 78.

prediction of individual observations and the evaluation of the influence of component variables on group membership.

### THE ANALYTIC TECHNIQUE

It is the objective of this research to classify a set of observations into mutually exclusive and exhaustive categories based upon a set of independent variables operating together. The discriminant function providing optimum classification of individuals into n number of discrete groups is of the following generalized form:<sup>3</sup>

$$Z_K = b_0 + \sum_{i=1}^j b_i X_i$$

where:

- $Z_K$  = is the value of the function in the  $K^{\text{th}}$  category
- $X_i$  = is the value of the  $i^{\text{th}}$  independent variable
- $b_i$  = is the value of the  $i^{\text{th}}$  discriminant coefficient
- $b_0$  = is the value of the function constant
- $j$  = is the number of independent variables.

#### Assumptions of the Model

Four assumptions are basic to the model.<sup>4</sup> The first assumes all groups are multivariate normal with mean vectors  $M_1$  and  $M_2$  and covariance matrices  $V_1$  and  $V_2$ . The second

<sup>3</sup>Donald G. Morrison, "On the Interpretation of Discriminant Analysis," Journal of Marketing Research, V6, (1969), p. 156.

<sup>4</sup>Ibid., pp. 162-63.

assumption is the covariance matrices  $V_1$  and  $V_2$  are equal. Third, it is assumed the estimated sample statistics are equal to the true population parameters and, lastly, the set of the explanatory variables are fully independent.

If it is assumed the original  $X_i$  are multivariate normal, then . . . "it becomes apparent that the discriminate function variate can be considered as having a normal distribution within groups".<sup>5</sup> Hence, all resulting linear functions are normal and the deviation of an individual discriminant score from its group mean can be transformed into a unit normal score.

The assumption of equal covariance matrices provides that the criteria or rule for classifying individuals into groups remains a linear rather than a quadratic function of the original variables.<sup>6</sup> Unequal covariance matrices increase or decrease the likelihood of an individual belonging to a certain group based upon the values of the previous  $X_i$ 's since the farther or nearer an individual  $X_i$  is to the common mean vector, the more likely this individual will belong to a certain group.<sup>7</sup>

---

<sup>5</sup>Overall and Klett, Applied Multivariate Analysis, (New York: McGraw-Hill, 1972), p. 246.

<sup>6</sup>Phoebus J. Dhrymes, Econometrics: Statistical Foundations and Application, (New York: Harper and Row, 1970), pp. 67-68.

<sup>7</sup>Morrison, op. cit., pp. 162-63.

set of Lastly, high correlations between the explanatory variables will alter the value of the discriminant coefficients and thereby confuse the interpretation of their effect on the total discriminant score.

The discriminant model provides one or more linear combinations (of variables) which have a maximum potential for discriminating among members of different groups by providing maximum average separation between the groups relative to within group variability.<sup>8</sup> In other words, the discriminant function maximizes the ratio of the variance between groups relative to the variance within groups. This is equivalent to saying the weighting coefficients are to be derived such that the t-statistic or F-ratio between groups will be maximum.<sup>9</sup>

Simply then, the discriminant problem amounts to choosing the b's and S's in such a way as to maximize the probability of correct classification. If the set of discriminant coefficients are optimum, the function will be maximized.

#### The Linear Classification Procedure

A set of linear functions of the independent variables have already been theorized to classify individuals or objects into discrete categories. It is necessary to discover from this set of functions a set of weighting coefficients for each

---

<sup>8</sup>T. W. Anderson, An Introduction to Multivariate Statistical Analysis, (New York: John Wiley and Sons, 1958), p. 137.

<sup>9</sup>Overall and Klett, op. cit., p. 244.

set of continuous vectors such that a boundary can be established between dichotomous groups.<sup>10</sup> The critical boundary values are determined by the number of independent variables, the boundary generally being an  $n-1$  dimensional hyper-plane in  $n$ -space.<sup>11</sup>

The classification rules are as follows:

(1) classify individual  $i$  as belonging to Group I if:

$$Z_i > Z_{crit}$$

(2) classify individual  $i$  as belonging to Group II if:

$$Z_i < Z_{crit}$$

$$\text{Given: } Z_i = b_0 + b_1X_{1i} + b_2X_{2i} + \dots + b_nX_{ni}$$

where:

$Z_{crit}$  = is the critical value for the discriminant score

$Z_i$  = is the  $i^{\text{th}}$  individual's discriminant score

$X_{ji}$  = is the  $i^{\text{th}}$  individual's value of the  $j^{\text{th}}$  independent variable

$b_j$  = is the discriminant coefficient for the  $j^{\text{th}}$  variable

$b_0$  = is a constant term in the discriminant equation

Since one assumption of the model is that of multivariate normality within groups, individual scores may be

<sup>10</sup>Judith M. Tanur, Frederick Mosteller, et al, Statistics: A Guide to the Unknown, (Holden Day Inc., 1972), pp. 234 pp. 234-36.

<sup>11</sup>Morrison, op. cit., p. 156.

transformed into unit normal scores. From unit normal distribution tables, cutting points may be chosen to adjust the number of misclassifications in any group.<sup>12</sup>

## APPLICATIONS OF THE DISCRIMINANT FUNCTION

### Determining Statistical Distance

The primary objective in utilizing a discriminant model is to identify some combination of variables that serves to separate  $n$  observations into  $m$  groups. The first question is whether the proposed function is significant in its ability to separate the groups; i.e., the function not only discriminates, but it discriminates with a specific level of precision. Although the answer to this particular question is deferred until the next chapter, the theoretical foundations for it have already been established.

Recall the discriminant function provides for maximum separation between group mean scores and minimum dispersion within each group. Since the difference or distance between group discriminant scores is based upon the same set of independent variables working together, it is reasonable to wonder whether this discriminant distance can be ascribed to chance

---

<sup>12</sup>For an extensive review of the theoretical basis for establishing optimal discriminant cutting points, refer to the following publications: David L. Sills, op. cit., pp. 554-57; C. I. Bliss, Statistics in Biology, (New York: McGraw-Hill, 1970), pp. 328-31; Maurice G. Kendall, A Course in Multivariate Analysis, (Charles Griffin and Company, 1965), pp. 145-47, 150-54; Overall and Klett, op. cit., pp. 247-48.

alone, and whether this distance or some measure of it is a reliable measure of the model's ability to discriminate.

A test statistic measuring the divergence of two populations was designed in 1921 by Karl Pearson.<sup>13</sup> Termed the "coefficient of racial likeness" (CRL), Pearson suggested the following form for the statistic:

$$\frac{N_1 N_2}{N_1 + N_2} (\bar{X}_1 - \bar{X}_2) S^{-1} (\bar{X}_1 - \bar{X}_2)$$

where:

$\bar{X}_i$  = the sample mean vector based upon a sample of size  $N_i$  from the  $i^{\text{th}}$  population.

$S$  = the pooled sample covariance matrix.

The dependent variate version of this statistic was modified between 1927 and 1930 by Morant and Mahalanobis to take the name of the  $D^2$  statistic and the generalized form of:

$$D^2 = \sum_{i=1}^m \sum_{j=1}^m \lambda_{ij} \sum_{k=1}^g N_k (\bar{X}_{ik} - \bar{X}_i) (\bar{X}_{jk} - \bar{X}_j)$$

where:

$i, j = 1, \dots, m$  are the socioeconomic variables

$k = 1, 2, \dots, g$  are the size of the operation groupings

$\lambda_{ij}$  = the  $i^{\text{th}}, j^{\text{th}}$  element of the pooled dispersion matrix

$N_k$  = sample size of the  $k^{\text{th}}$  group

<sup>13</sup>Das Gupta, op. cit., pp. 78-79.

$\bar{X}_{ik}$  = mean of the  $i^{\text{th}}$  variable in the  $k^{\text{th}}$  group

$\bar{X}_i$  = the overall mean of the  $i^{\text{th}}$  variable<sup>14</sup>

The  $D^2$  statistic represents a generalized distance between two groups with the same set of variables and identical variance-covariance matrix. Under the assumption of independence of the explanatory variables, the  $D^2$  statistic can be characterized by its  $n$  dimensional mean vector. It is simply the square of the usual Euclidian distance between two mean vectors, where the orthogonal co-ordinate system is normalized by the standard deviation of each variable.<sup>15</sup>

The desirable properties of decision rules based upon the  $D^2$  statistic are due to the fact that it "emerges as the natural measure of dissimilarity between homoscedastic normal populations".<sup>16</sup> Because the desirable properties of this statistic are consistent with the underlying assumptions of the model and can serve to evaluate the significance of group distance, the evaluation of the model's ability to discriminate will be based upon this statistical measurement.

<sup>14</sup>A.A. Araji and R.M. Finley, "Managerial Socioeconomic Characteristics and size of Operation in Beef Cattle Feeding: An Application of Discriminant Analysis," American Journal of Agricultural Economics, LIII, 4, (1971), 648.

<sup>15</sup>Morrison, op. cit., p. 157.

<sup>16</sup>T. Cacoullos, "Distance, Discrimination, and Error," in Discriminant Analysis and Applications, (New York: Academic Press, 1973), pp. 61-67.



### Evaluating the Variables

A second objective in using discriminant analysis is in evaluating the influence of individual explanatory variables upon the total individual discriminant score. The general form of the discriminant model reveals the importance of the explanatory variables to the individual discriminant score. Because the classification of an individual ultimately depends upon the value of his discriminant score, the contribution of the discriminant coefficients and the explanatory variables should offer a foundation for evaluating and ranking the influence of any one particular variable. Chapter 3 will examine several methods of evaluating the influence of a specific variable based upon different measures of its contribution to the discriminant score.

### Classifying Individuals into Groups

Lastly, the discriminant function serves to classify individuals into discrete categories. However, individuals may or may not be assigned to the discrete category which they have been assigned, a priori. This possibility of misclassification in itself suggests a measure for evaluating the efficacy of the model: the ability to successfully classify individuals on a greater than chance basis. Chapter 3 also contains a discussion of the criteria for evaluating the predictability of the model.

Thus, a discriminant function allows the classification of individuals through maximum possible separation of group

discriminant scores based upon a linear combination of variables. From this, three applications of discriminant analysis have been suggested for examination as evidence of significance in the model: they are (1) the ability to test the statistical significance of the distance between group discriminant scores, (2) the ability to evaluate and rank the impact of individual explanatory variables, and (3) the ability to successfully classify individuals into groups. Examined in Chapter 3 are all three suggested measures of significance in light of their contribution to interpreting fundamental group differences.

#### THE WELFARE POPULATION

The general welfare population which this study examines is ADC recipients within the State of South Dakota. This particular group of welfare recipients was chosen for several reasons. First, almost three of every four welfare cases in South Dakota fall under the ADC program while the remaining fourth are divided between Old Age Assistance, Aid to the Blind, and Aid to the Disabled.<sup>17</sup>

Secondly, welfare programs other than ADC are comprised of individuals with serious obstacles to gaining economic self-sufficiency. The blind, disabled, and the aged are forced,

---

<sup>17</sup> January 1973 study of 777 ADC recipients by the Department of Social Services, South Dakota.

almost entirely through circumstances beyond their control, to accept their position as dependents of the state. ADC recipients are more fortunate in that a range of programs are available to allow individuals some possibility for employment, additional income, and a chance to escape the stigma of being a welfare recipient.

### Delimiting the ADC Population

The specific population of ADC recipients used in this study has been limited to unmarried ADC mothers in Pennington County, and is further restricted to include only recipients from the geographic confines of Rapid City, South Dakota. The reason for these restrictions are three-fold. First, ADC information was most easily obtained through the county and city because of the presence of a regional ADC office within Rapid City. Secondly given the large number of ADC recipients within Pennington County (approximately 3,000) most residing within Rapid City, a representative sample could be obtained with a minimum expenditure of time and money.<sup>18</sup> Thirdly, the population was restricted to unmarried recipients in order to eliminate the influence of the husband's income upon income expectations.

The list of ADC recipients was obtained from monthly payment listings through the Pennington County welfare office.

---

<sup>18</sup>South Dakota Department of Social Services, September 1973 Monthly PAR report.

The list, containing all active ADC cases for the month, is formatted so those recipients with the longest duration in the ADC program appear first while the newest cases appear last. To achieve a representative cross-section of welfare families by length of stay in ADC, the list was divided into four equal clusters. From each cluster thirteen families were selected through the application of random number tables to each cluster for a total of 52 families, 31 displaying positive income expectations and 21 with negative expectations. The fifty-two families represent a five percent sample of welfare families in Pennington County and a fairly representative cross section of ADC recipients in South Dakota. An examination of some socio-economic characteristics of the South Dakota welfare population should offer some insight into factors possibly affecting income expectations.<sup>19</sup>

The ADC population in South Dakota at the time the sample was taken consisted of approximately 22,500 recipients, 43 percent of whom are of American Indian descent. This fact is of particular significance since nearly one Indian in three is an ADC recipient, yet this group comprises only five percent of the state population. Another interesting fact is that only one ADC mother in eight is employed full time (35 or more hours a week). An additional seven percent are employed on a part

---

<sup>19</sup>Study of ADC recipients, op. cit., Tables 3, 13, 14, 15, 16, 17, 27. All statistics in this section are from this study unless specified otherwise.

time basis. Of the remaining households, only 12 percent are actively seeking work and nearly 31 percent have never been gainfully employed. Even for ADC mothers fortunate enough to be fully employed, Census Bureau statistics reveal female household heads earn on the average only \$3,000 annually.<sup>20</sup> Finally, a fourth of all active ADC recipients have been receiving assistance for over four years.

The sample information used in this survey was obtained through telephone interviews during the month of August, 1973, with all 52 families. The primary advantages of this technique are the convenience, speed, and efficiency with which telephone interviews may be conducted. Two major drawbacks do exist with telephone sampling, especially in terms of this study.

Telephone sampling may introduce an element of economic bias by avoiding families who do not have or cannot afford telephone service. This particular omission is likely to be more noticeable when sampling is within the lowest income groups, as is the case in the present study.

In addition, more than one third of the families drawn from the random sample were unable to be contacted for a number of reasons, including bogus telephone numbers, disconnected phone service, or new or unlisted phone numbers.

---

<sup>20</sup>1970 Bureau of the Census data from Minnesota Data and Analysis Planning System, p. 1, Table 77.

Secondly, an element of bias may be introduced in telephone surveys since those individuals interviewed by telephone may be less candid than if interviewed personally.

The specific survey questions are contained in Appendix I. The questions are almost entirely of two basic formats which offer the respondent either a dichotomous response choice or a scaled response choice. The following are examples of each, respectively.

Are you presently employed?

Yes

No

How strongly do you feel your income will increase (decrease) in the next twelve months?

1

2

3

4

5

6

7

Decrease  
strongly

Neutral

Increase  
strongly

For purposes of scaling, the dichotomous questions have been adjusted to the same numeric scale as the scaled response questions. The reasons for this adjustment are discussed below.

The use of dichotomous questions coupled with supporting questions of the scaled variety allows the respondent to categorize himself into one of two groups which simplifies his problem of choice. Next, scale response in supporting questions allows a full but non-complex range of alternatives from the very strong to very weak to neutral responses regarding a question. Thus, the two types of questions allow a full

but unobtrusive range of responses that are easily quantified for later interpretation.

### THE CHOICE OF VARIABLES

The basis for evaluating income expectations is subject to a wide range of beliefs. Many investigators contend expectations are purely psychological considerations subject to an infinite variety of influences. They would assert the analysis of such subjective considerations is not possible since there are far too many factors influencing expectations and this set of influences is constantly changing. This study prefers to take the position that although expectations are difficult to identify, they are not beyond the realm of identification and quantification.

The ten variables selected for examination in this study fall into three general classes. The first class of variables relates to the individual recipient's present and past income and the ability and perceived ability to maintain or gain employment opportunities. The second class of variables relates expectations to borrowing and saving and the perceived ability to borrow or save. Additionally, the recipient's self-perceived ability to leave ADC is also examined. The tenth and final variable relates to the racial origin of the welfare recipient. A brief discussion of the reasons for including these variables is outlined below.

Concerning the first group of variables, annual gross estimated income has been included since it is the most

concrete measure of an individual's economic success relative to other welfare recipients and the rest of the income distribution. It is suggested this relative position in the welfare distribution as well as one's absolute level of income has a influence upon future income expectations and ultimately serves as a basis for comparison in future years.

Peak income is included as a variable for the reason just mentioned. It is possible an individual who is above or below their previous peak income will perceive future income expectations differently due to a change in their relative standard of living.

Employment is chosen as an explanatory variable for several reasons. First, it would be illogical to voluntarily seek employment without the expectation of increasing one's income. Employment is basically the only avenue for an ADC recipient to substantially increase her monthly benefits and income.

The effect of employment on income expectations works in several ways. Employment yields more than one component of future income to the welfare mother. The recipient's net employment income is based on the state's rules for computing ADC benefits. States must allow the recipient the first thirty dollars and one third of the remaining employment earnings in order to provide a work incentive. This component of work income may actually be less important than the additional income received through deductions for working



expenses.<sup>21</sup> Most states included the cost of transportation, day-care for dependents, and special outlays for other work-related expenses such as special uniforms, union dues, lunches, and even income and social security taxes.<sup>22</sup> In addition to the work related benefits, the working ADC mother may also expect to receive income benefits through her employer in the form of pensions, extra medical benefits, and cheaper credit through employee credit unions.

The employment expectations variable is an index of the recipient's belief in the availability of full time employment opportunities. The expectation of continuing employment for those recipients presently employed or expecting employment should contribute to a set of income expectations different from the expectations of recipients who do not perceive the ability to gain employment.

The second group of variables relates to expectations or economic behavior that should differ between groups with different income expectations. Since conventional economic theory suggests the short run marginal propensity to save is an increasing function of income, it is possible that saving expectations would be related to income expectations in some

---

<sup>21</sup>The working ADC mother may receive as much as \$189 dollars extra each month in child care payments (\$129) and related payments (\$60). Additional support payments are available as well.

<sup>22</sup>Irene Lurie, "Estimates of Tax Rates in the AFDC Program," National Tax Journal, XXVII, 7, (1974), 93-106.

manner. By a similar line of reasoning current savings of those individuals with positive income expectations should be different from individuals exhibiting negative expectations.

The perceived ability to borrow and the desire to borrow are suggested as related to income expectations through the increasing availability of credit as income rises. Rising income also increases the possibility of a reduced cost in obtaining credit by qualifying for lower interest rate loans, perhaps through commercial banks. Thus, individuals with positive income expectations and the availability of credit may be more likely to forego future consumption for current consumption.

Finally, the perceived ability to leave ADC is usually predicted on the belief of a much higher level of future income. This change in income is usually through a change in marital status, or a fortuitous change in employment. The departure of children from the household could also contribute to higher per capita income. Since all of these possibilities will likely result in a change in future income, groups with different income expectations should be expected to exhibit different perceived abilities to leave ADC unless other factors exist which discourage a recipient from leaving ADC.

The tenth variable identifies the racial characteristics of those recipients contacted in this study. For the purpose of this investigation, only two racial groups appeared, Caucasian and American Indian.

The inclusion of the racial variable could be related to income expectations in several ways. Different ethnic groups appear to have unequal opportunities in receiving proper education or gaining access to financially rewarding employment. Certain ethnic groups may also lack motivations or expectations based upon a heritage of subservience and poverty. Moreover, certain ethnic groups appear to be more disposed to accepting poverty status based upon a history of economic suffering. It is an unfortunate fact that the results of this study cannot confirm or deny these influences. The study only attempts to answer the question of whether race or any other variable affects income expectations, not why. With this fact in mind, the results of the study will now be presented.

## Chapter 3

### ANALYSIS OF THE RESULTS

In discussing the theoretical foundations of the model, three measures of significance were suggested to help in evaluating the model. These measures are (1) the level of significance at which the model can discriminate between groups, (2) the proportion of correct group classifications, and (3) the relative contribution of each explanatory variable to the total discriminant score.

The analysis of results presented in this chapter are based largely upon these three measures of significance. Each of these areas is examined in light of their contribution to the interpretation of just how the model helps discriminate between groups and to the interpretation of fundamental group differences.

### RELEVANCE OF VARIABLES TO THE MODEL

The choice of variables in a statistical model is, to a certain extent, at the discretion of the researcher. Most variables are chosen because they have previously been identified as causally related to the effect being explained, or the researcher is hypothesizing such a relationship. However, the researcher may add variables which offer little or nothing to the explanatory power of the model, or may omit variables offering significant explanatory power.

In the latter case, the omission is likely due to an incomplete understanding of all factors influencing the model. The researcher is simply unaware or is unable to identify all variables relevant to the model. Omissions of this type are often easy to identify, but difficult to remedy. In the former case the research may include one or more variables relevant to the model but having little power to discriminate. It is also possible to include variables appearing to have real explanatory power that are very closely associated with the effects of some other variable.

Variables relevant to the model but making little contribution to the total function are retained or discarded based upon the researcher's criterion for significance and the desired level of precision. Several methods for evaluating the relative importance of explanatory variables are presented later in this chapter.

In the case of highly correlated explanatory variables, the researcher is required to either remove or adjust the variables since this violates a major assumption of the model: independence of the explanatory variables. If this adjustment or removal is not undertaken, the value of the coefficients attached to each variable will be altered and the interpretation of the relative importance of each variable will be obscured.

A measure of association between variables has been listed in Table 3-1. This measure of association is called

Table 3-1

Matrix of Correlation Between  
Discriminant Variables

| Variable | 2    | 3    | 4    | 5    | 6   | 7    | 8    | 9   | 10   |
|----------|------|------|------|------|-----|------|------|-----|------|
| 1        | -.15 | .50  | .61  | .04  | .16 | .33  | .43  | .14 | .01  |
| 2        |      | -.03 | -.03 | -.08 | .27 | .001 | -.15 | .22 | .06  |
| 3        |      |      | .58  | -.25 | .05 | .33  | .31  | .17 | .09  |
| 4        |      |      |      | -.04 | .10 | .28  | .44  | .17 | .14  |
| 5        |      |      |      |      | .25 | -.07 | -.05 | .03 | -.22 |
| 6        |      |      |      |      |     | .17  | .20  | .68 | .03  |
| 7        |      |      |      |      |     |      | .37  | .25 | .09  |
| 8        |      |      |      |      |     |      |      | .40 | .23  |
| 9        |      |      |      |      |     |      |      |     | .19  |

the "correlation coefficient" and it relates the joint movement or association between variables.<sup>1</sup> Examination of Table 3-1 shows no evidence of exceptionally strong associations between variables. The absence of strong correlations between variables tends to support the assumption that the movement of variables used in this model are relatively independent.

#### STATISTICAL SIGNIFICANCE OF THE MODEL

The question of statistical significance is based upon the properties of the linear discriminant function which provide for maximum separation between group discriminant scores and minimum dispersion within groups. It is the property of maximum separation that will serve as the basis for our test of statistical significance.

Chapter 2 suggested a measure of statistical significance could be established through testing the generalized distance between groups based upon the assumption this distance was attributable to chance alone. The generalized Mahalanobis D-Square statistic was presented as a measure of this distance between groups. The D-Square statistic is distributed as a chi-square statistic with  $N(G-1)$  degrees of

---

<sup>1</sup>William Mendenhall, An Introduction to Probability and Statistics, (Wadsworth Publishing Co., 1967), pp. 240-42.

freedom, where  $N$  is the number of variables and  $G$  the number of groups.<sup>2</sup>

The value of the D-Square statistic computed in this study is 42.83426, which is significant at the .001 level with ten ( $10(2-1)$ ) degrees of freedom. A high level of significance indicates the ability to discriminate between groups on the basis of the combined effect of the variables. In other words, the function can successfully discriminate between group membership is considered next.

#### CLASSIFICATION OF INDIVIDUALS BY GROUP

A second measure of significance of the model is its ability to correctly classify individuals. One approach to this problem is the construction of an  $n \times n$  classification matrix found in Figure 3-1 below.<sup>3</sup> The matrix plots actual group membership against predicted group membership.

---

<sup>2</sup>A.A. Araji and R.M. Finley, "Managerial Socioeconomic Characteristics and Size of Operation in Beef Cattle Feeding: An Application of Discriminant Analysis," American Journal of Agricultural Economics, LIII, 4, (1971), 648.

<sup>3</sup>Donald G. Morrison, "On the Interpretation of Discriminant Analysis," Journal of Marketing Research, V6, (1969), 156.



|          |         | Classified    |               |              |
|----------|---------|---------------|---------------|--------------|
|          |         | Group 1       | Group 2       |              |
| Actually | Group 1 | $n_{11}$      | $n_{12}$      | $n_{1\cdot}$ |
|          | Group 2 | $n_{21}$      | $n_{22}$      | $n_{2\cdot}$ |
|          |         | $n_{\cdot 1}$ | $n_{\cdot 2}$ | $n$          |

Figure 3-1

The proportion of individuals correctly classified is the ratio of  $(N_{11} + N_{22}) / N$ . The confusion matrix constructed for this study is found in Figure 3-2. An application of the confusion matrix is the computation of a Q-statistic. The Q-statistic is distributed as a chi-square statistic with one degree of freedom and has a value of 24.923, significant at the .001 level.<sup>4</sup>

|        |         | Confusion Matrix |         |
|--------|---------|------------------|---------|
|        |         | Group 1          | Group 2 |
| Actual | Group 1 | 27               | 4       |
|        | Group 2 | 4                | 17      |
|        |         | Classified       |         |

Figure 3-2

<sup>4</sup>James Press, Applied Multivariate Analysis, (New York: Holt, Rinehart and Winston, 1972), pp. 381-82.

However, both the Q-statistic and the confusion matrix are subject to a built-in upward bias since the observations used to calculate the discriminant function are the same observations classified by these two procedures. A slight adjustment to the results of each procedure is generally required.<sup>5</sup>

The confusion matrix approach shows the ability of the model to predict total group membership with an 84.6 percent accuracy. An examination of Figure 3-2 shows similar predictability with respect to groups individually.

These figures suggest a reasonable level of predictive ability can be ascribed to the model. But, since both groups are not of equal size, the chance probability of drawing an individual from either group is unequal. Given groups of unequal size, the chance model for group classification may be specified.<sup>6</sup> The chance model applicable to the present study yields only slightly better than 51 percent change of random classification into Group I, as shown in Table 3-2. Hence, our model offers a reasonably good level of predictability.

---

<sup>5</sup>Morrison, op. cit., p. 157.

<sup>6</sup>Ibid, pp. 158-60.

Table 3-2

## Misclassification Probabilities

$$C_{\text{pro}} = .5184$$

$$C_{\text{max}} = .596$$

where:

$$C_{\text{pro}} = p \alpha + (1-p)(1-\alpha)$$

$p$  = true proportion of Group I individuals

$\alpha$  = proportion of individuals classified  
as Group I

## EVALUATING THE VARIABLES

Assessing the importance of the variables is to a large degree determined by the size of the discriminant coefficients since they determine the total absolute contribution of an individual variable to the total discriminant score. Those coefficients with large numerical values make the largest unadjusted numerical contribution to the discriminant score. Hence, the absolute value of the discriminant coefficients will be considered as one measure of the relative importance of the explanatory variables.

Another related method of evaluating the discriminant variables is to assess the contribution of the mean value of each variables times its discriminant coefficient. This provides a proxy of the average contribution of each variable to the total discriminant score. In this manner, the explanatory variables may be ranked according to their average

contribution to the discriminant function. This technique will also be considered as one possibility for ranking the explanatory variables.

Both of these methods provide a basis for evaluating the relative contribution of all the variables, but each suffers from the same serious deficiency. Disregarding sign, a relatively large discriminant coefficient does not necessarily indicate a measure of greater importance than one with a smaller coefficient. The reason is the value of the coefficients will vary in magnitude with the number of individuals at each  $X_i$  and with their scaling.<sup>7</sup>

To adjust for these deficiencies, the coefficients may be set to the same scale or "normalized" by (1) dividing the product of the coefficients and their respective group means by the group standard deviation, or (2) subtract the group standard deviation from its group mean and multiply this by the respective discriminant coefficient. Both adjustments account for differences in scale and the second has the added advantage of differentiating between variables with identical standard deviations.<sup>8</sup> This last suggested transformation is used as the basis for examining the importance of each

---

<sup>7</sup>C.I. Bliss, Statistics in Biology, (New York: McGraw-Hill, 1970), p. 335.

<sup>8</sup>James R. Prescott and William C. Lewis, "State and Municipal Locational Incentives: "A Discriminant Analysis," National Tax Journal, XXII, 3.

variable in this study. A ranking of the explanatory variables using all of the suggested methods is found in Table 3-3.

### Evaluating the Income Variable

An examination of Table 3-3 reveals that, regardless of the method used to evaluate the variables, estimated gross annual income is clearly the most important variable in terms of its ability to discriminate between groups. However, the relative contribution of this variable to each group's discriminant score is not the same. This fact suggests some possible differences between the two groups.

For the most part, families of both groups are of similar family size, similar age, and of similar educational background. All are unmarried and all are ADC recipients. Based on this information, each group should have approximately the same average monthly ADC payment, barring work income. Yet, individuals of Group I are most effectively distinguished from their Group II counterparts by income.

Since this difference in income cannot be accounted for by large group differences in ADC payments, another explanation must be found. The alternative explanation coming first to mind is in differences in the two employment variables. Examination of row 4, Table 3-3, indicates these two employment variables are of considerable importance in the model. The first employment variable, re-employment expectations, refers to the self-perceived ability to maintain

Table 3-3

## Ranking of Discriminant Coefficients

| Variable                     | 1 | 2 | 3 | 4 | 5 | 6  | 7  | 8 | 9  | 10 | Group |
|------------------------------|---|---|---|---|---|----|----|---|----|----|-------|
| $K_i$                        | 1 | 3 | 2 | 6 | 4 | 7  | 10 | 9 | 8  | 5  | I     |
|                              | 1 | 2 | 4 | 3 | 5 | 7  | 9  | 8 | 10 | 6  | II    |
| $K_i \bar{X}_i$              | 1 | 3 | 2 | 5 | 4 | 7  | 10 | 9 | 8  | 6  | I     |
|                              | 1 | 2 | 3 | 5 | 4 | 6  | 9  | 8 | 10 | 7  | II    |
| $\frac{K_i \bar{X}_i}{Sx_i}$ | 1 | 3 | 2 | 5 | 4 | 7  | 10 | 9 | 8  | 6  | I     |
|                              | 1 | 2 | 3 | 5 | 6 | 7  | 10 | 9 | 8  | 4  | II    |
| $K_i(X_i - Sx_i)$            | 1 | 3 | 2 | 4 | 5 | 10 | 9  | 8 | 7  | 6  | I     |
|                              | 1 | 2 | 4 | 3 | 7 | 6  | 9  | 8 | 10 | 5  | II    |

## Variables

$K_i$  = Value of the  $K^{\text{th}}$  discriminant coefficient

$\bar{X}_i$  = Value of the  $X^{\text{th}}$  discriminant variable (mean)

$Sx_i$  = Value of Group discriminant standard deviation

- |                              |                              |
|------------------------------|------------------------------|
| (1) Est. Gross Annual Income | (6) Savings                  |
| (2) Race                     | (7) Desire for Credit        |
| (3) Employment Expectations  | (8) Ability to Get Credit    |
| (4) Employment Status        | (9) Ability to Save (Future) |
| (5) Peak Income              | (10) Ability to Leave ADC    |

employment for those recipients presently employed, and the self-perceived ability to gain full time employment for those recipients currently unemployed. The second employment "variable" is whether the recipient is currently employed or unemployed.

In Group I a much greater percentage of recipients were employed than in Group II (58 percent vs. 14 percent). Perhaps even more significant are the implications that may be drawn from the importance of the employment expectations variable.

First, we have evidenced the fact that a much larger proportion of Group I recipients are employed. This difference in employment can account for much of the difference in income between the two groups. Perhaps as important is the fact that Group I has a very strong perception of its ability to maintain or gain employment. So important are employment expectations to Group I that they rank as the second most important variable. This variable is only the fourth best discriminator in Group II, its contribution being much less in Group II than in Group I.

The distinction may be found in the large difference between each group in the way employment opportunities are perceived. Group I, with the largest number of employed and with the highest incomes, represents the "wealthy" class of welfare recipients; a class of recipients with a much larger degree of financial freedom and a class of recipients with a

potential opportunity to completely escape from the welfare system by virtue of their employment.

Group II, being largely unemployed, is resigned to total financial dependence on ADC payments. For those individuals in Group II who are largely discouraged about their opportunity for re-employment, the only hope for increasing their income is through changes in the basic payment formula.

Members of Group I who are employed or who strongly perceive they will be employed can expect some nominal increase in their wage rate in addition to any change in their ADC payments. This group is more likely to have a recognizable increase in their nominal money income from year to year; enough perhaps to have a decided impact upon future income expectations.

#### Evaluating Non-Income Variables

Groups I and II differ fundamentally in terms of their estimated gross income and in the relative success of each group in gaining employment. It has been suggested that an important clue to understanding the importance of the remaining variables is the rather striking difference in the position of the two employment variables.

The decline of employment expectations from the second best discriminatory variables in Group I to the fourth best discriminatory variable in Group II suggests a shift in importance of other variables.



A measure of the significance of this movement may be found in the contribution of these variables, given the contribution of income. In other words, the importance of the other nine variables to the discriminant score, holding the effect of income as a constant, may offer some added insight into our problem.

Table 3-4 contains the percentage contribution of all ten variables to the discriminant score as well as the contribution to the discriminant score of the nine variables after the impact of income has been considered.

The fact that income makes a much greater contribution to the discriminant score in Group II (52 percent to 31 percent) implies the residual to be explained in Group I is greater than in Group II. The contribution of the remaining variables can offer a basis for some inferences.

Group I is distinguished by its positive attitude toward re-employment. The contribution of this variable to the total discriminant score is over one fourth of all variables combined while the same variable in Group II contributes less than one twelfth of the total discriminant score. Even after considering the impact of income, employment expectations contribute well over two times as much to the residual in Group I as in Group II.

The impact of the employment status variable is just the opposite. In terms of this variable's contribution to the total function, Group II is more affected by unemployment in

Table 3-4

Percent Contribution of Adjusted Discriminant  
Variables to Total Discriminant Score

| Variable*  | 1     | 2     | 3     | 4     | 5     | 6    | 7    | 8    | 9    | 10    | Group |
|--|-------|-------|-------|-------|-------|------|------|------|------|-------|-------|
|  | 31.16 | 21.24 | 25.46 | 7.24  | 7.15  | 0.29 | 0.47 | 1.72 | 1.87 | 3.41  | I     |
|  | 52.44 | 17.23 | 7.97  | 9.37  | 2.23  | 3.85 | 0.56 | 1.11 | 0.24 | 5.00  | II    |
| Percent Contribution of Adjusted Discriminant<br>Variables With Estimated Gross Annual<br>Income Held Constant |       |       |       |       |       |      |      |      |      |       |       |
|  |       | 30.85 | 36.98 | 10.52 | 10.38 | 0.42 | 0.68 | 2.50 | 2.72 | 4.95  | I     |
|  |       | 36.23 | 16.76 | 19.70 | 4.69  | 8.10 | 1.18 | 2.33 | 0.50 | 10.51 | II    |

\*See Table 3-3 for listing of variables

terms of discriminatory power than is Group I. This difference is even more pronounced when income is held constant. Employment status explains nearly two times as much of the residual in Group II as in Group I.

These two comparisons highlight the fact that Group I recipients are largely working or expect to get work soon. This accounts for the importance of employment expectations and the lesser importance of employment status. On the other hand, Group II is largely unemployed and almost lacks expectations of work. Since unemployment in Group II is so prevalent, this variable has an increased importance in terms of the discriminant function. However, another factor remains to be considered. This remaining factor is race.

The present study examined only Caucasian and American Indian families. While less than one in five recipients from Group I were Indian, almost half of Group II were Indian. This particular fact offers some additional evidence to support the contention that it is possible to discriminate between our two groups on the basis of certain observed characteristics. It also contributes to the understanding of the relative changes in importance of other variables.

By our earlier methods of evaluation, race is the third most important variable in Group I and the second most important variable in Group II. Examination of Table 3-4 indicates this variable is important not only to the total function but

to the residual after the effects of income have been considered.

The numerical contribution of this variable is important in each group and its magnitude for each is much the same. However, the implication is not that race works in the same direction in discriminating between groups. Rather, for each group, race is the opposite side of the coin. Group I, largely Caucasian, displays positive income expectations, positive employment expectations, has the highest income, and is apparently more successful in obtaining full time employment. Group II, with a much larger proportion of Indian families, fails to exhibit any of the above tendencies.

It would not be fair to conclude these differences are due entirely to the presence or absence of a single variable. However, due to the large discriminatory power of the racial variable, it would be difficult to conclude that race does not contribute to these differences.

Group I has a selective advantage over Group II in several categories, and it is difficult to identify one single reason for these differences. The attempt has been made to point out the relationship between those variables contributing most to the separation of our two groups and gain some insight into how the variables work to determine the level of individual income expectations. The story is by no means complete and at least one additional index of group differences remains to be examined.

### The Desire to Leave ADC

An examination of Table 3-4 leaves the impression that variable ten, the recipients perceived ability to leave ADC, offers little to the discriminant function's ability to separate our groups. The contribution of this variable to the total discriminant score of each group is never over five percent of the total. Examination of Table 3-5 also shows that while a difference exists in the group mean scores for this variable, each score reflects a very negative assessment of the ability to leave ADC. Since both group's set of responses are much the same, this variable does not provide a statistical basis for reliable group separation. However, important implications are suggested by this fact.

A likely explanation of the almost non-existent desire to leave is found in the opportunity cost of abandoning ADC. The cost is this; a family earning one dollar less than the maximum amount allowed in order to remain on ADC is still entitled to receive full medical coverage for the recipient's family.

It is clear the cost of abandoning ADC is great. In dollar terms, this cost would be nearly equal to the cost of full comprehensive medical insurance. It is unlikely this cost could be met even with a very profitable change in employment. Consequently, an upward limit on employment and job income is placed on the ADC recipient who is earning at or near the maximum income allowed under existing guidelines.

Table 3-5

Value of Group Means, Discriminant Coefficients, and  $D^2$  Statistic

---



---

| Mean Values                             |        |        |        |         |        |         |         |        |        |        |       |
|---|--------|--------|--------|---------|--------|---------|---------|--------|--------|--------|-------|
| Variable*                               | 1      | 2      | 3      | 4       | 5      | 6       | 7       | 8      | 9      | 10     | Group |
|   | 4.5484 | 5.8387 | 5.9032 | 4.2903  | 4.2903 | 2.3548  | 3.4839  | 4.0323 | 3.5806 | 2.6774 | I     |
|   | 3.6190 | 4.1429 | 3.1905 | 1.8571  | 3.2857 | 1.5714  | 1.7143  | 2.2381 | 1.6667 | 1.1429 | II    |
| Discriminant Coefficients               |        |        |        |         |        |         |         |        |        |        |       |
| Constant -12.97741                      |        |        |        |         |        |         |         |        |        |        |       |
| Variable*                               | 1      | 2      | 3      | 4       | 5      | 6       | 7       | 8      | 9      | 10     | Group |
|   | 1.8462 | 1.1780 | 1.2259 | -.80228 | .97975 | -.69224 | .06830  | .22480 | .26820 | .82425 | I     |
| Constant -6.85657                       |        |        |        |         |        |         |         |        |        |        |       |
|   | 1.9715 | .90041 | .66322 | -.82632 | .62768 | -.41112 | -.07650 | .18123 | .02595 | .46580 | II    |
| Generalized Mahalanobis $D^2$ Statistic |        |        |        |         |        |         |         |        |        |        |       |
| 42.82436                                |        |        |        |         |        |         |         |        |        |        |       |

---

\*See Table 3-3 for listing of variables

expect The contribution of the remaining five variables constitutes less than ten percent of the total discriminant score for each group. As with the case of the ability to leave ADC, the responses to the remaining five categories were close enough to reduce the basis for discriminating between groups with these particular variables.

strong What remains is the evidence that tends to conclude the model actually can discriminate between groups based on the higher percentage of correct classification, the significance of our D-Square statistic, and the ability of the model to provide a basis for identifying those variables most important to the discriminant function.

groups In conclusion, each group shares the same set of four variables having a large discriminatory power. The relative importance of these variables highlights the difference between each group.

First, estimated gross annual income contributes the most to the discriminant score of each group: Group I receiving the highest annual income and Group II receiving the lowest. Group I is also contrasted with Group II through its strong positive employment expectations variable; the second best discriminator in this group. Additionally, Group I is blessed with a much higher level of employment.

Group II receives the lowest estimated gross annual income and suffers from a largely negative set of employment expectations. The decline in importance of the employment

expectations variable to the weakest of the four major variables emphasizes the inability of this group to obtain employment.

Race offers a final contrast between groups. Group I, the group enjoying positive income expectations, is largely Caucasian. Group II, evenly divided ethnically, suffers from strong negative income expectations.

Based largely upon these four variables, the discriminant model has identified significant differences between each group. These differences have been translated into a discriminant model which, on the basis of group discriminant distance, is highly significant in its ability to discriminate between groups, correctly classify individual observations, and provide a basis for evaluating the variables used in the model. With this in mind, the question of the achievement of this study's objectives is now addressed.



## Chapter 4

### SUMMARY, IMPLICATIONS, AND RECOMMENDATIONS

The purpose of the present study has been to better describe the elements underlying the income expectations of a subset of the general welfare population. The description has required three objective considerations: (1) the description of an underlying set of explanatory variables; (2) the analytic and statistical description of differences between sample groups; and (3) the description of the relative explanatory value of the underlying variables.

In each case, the study has been able to shed light on the foundations of income expectations. The study has identified a linear combination of ten variables that have been highly significant in discriminating between sample groups. In addition, four variables, estimated gross annual income, employment expectations, race, and employment status have been identified as those variables contributing most to the discriminant function. Thus, from the point of view of the researcher, the objectives of the study have been satisfactorily addressed.

#### IMPLICATIONS OF THE STUDY

Earlier in this study, one very important disincentive to achieving economic independence was identified. This

disincentive is the comprehensive medical coverage offered to ADC recipients. Since this research has been undertaken, the medical benefits available to ADC recipients have been expanded to include comprehensive dental care and medical prescriptions. Thus the monetary disincentive to depart the ADC program has grown.

The present study has highlighted the absence in either sample group of a perceived ability to leave ADC and has suggested the opportunity cost of abandoning these benefits as a major reason for remaining in the ADC program. Another important disincentive to leaving the ADC program lies in the range of employment opportunities available to the individual recipient.

In both sample groups, many families expressed the desire to find meaningful and rewarding employment but were unable to do so. Most recipients with jobs worked as low skilled employees and it appears the working welfare mother suffers from the double stigma of limited employment opportunities and limited financial rewards.

Both of the aforementioned areas should provide numerous topics for future research. Another topic demanding additional examination is the foundations of the apparent differences in income expectations exhibited by recipients of different racial origin.

The American Indian welfare recipient displays an entire range of negative expectations covering income,

employment, the availability of credit and so on. With nearly one third of the state's thirty thousand Indians supported through ADC, every effort should be undertaken to identify the fundamental reasons for this dependence upon public assistance. Moreover, unless attempts are made to encourage economic self sufficiency and provide means for reducing the number of Indians receiving welfare, State and Federal agencies can be assured of many additional generations of Indian poor.

#### RECOMMENDATIONS FOR FURTHER RESEARCH

One area related to the field of social welfare economics which warrants additional research is in the relationship between income expectations and the existing rules governing ADC payments. Closely related to this particular topic is an examination of the marginal tax rate on earned income and the marginal benefits of employment as an ADC recipient.

By Federal law, states are required, as a minimum, to allow the working recipient to keep the first thirty dollars of earned income and one third of the remainder. Deductions allowed in excess of this "thirty and a third" varies widely from state to state. An examination of differential rates of employment based upon work incentives would be useful in providing evidence to evaluate the importance of employment incentives on total employment.

Another suggested avenue for research would be another study of ADC recipients in Rapid City. With various changes

in the basic payment formula and in the availability of medical benefits, research should center on any change in the relative importance of the explanatory variables used in this study. Changes in the value of the discriminant coefficients should be examined in view of these changes.

Another area with very important normative implications is in the area of the opportunity cost of departing ADC. It is conceivable that a significant number of ADC recipients are maintaining an underemployed status so as to remain recipients of medical benefits.

Lastly, research should evaluate the ability of an individual recipient to leave ADC given the numerous possible disincentives to do so. Research should focus on families leaving for reasons other than a change in marital status. This research should focus on the likelihood of a recipient leaving ADC and returning to the program at some later date.

In conclusion, this study has focused on only one of hundreds of topics related to social welfare economics. It is hoped some small contribution has been made to the understanding of the economic foundations of income expectations and this knowledge can be applied to the improvement of the economic existence of the poor.

BIBLIOGRAPHY

Ackoff, Russell. Scientific Method: Orbits and Applications. New York: Wiley, 1962.

Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: Wiley, 1958.

Bliss, C. I. Statistics in Biology. New York: McGraw-Hill, 1952.

Cacoullos, T. Biometric Analysis and Applications. New York: McGraw-Hill, 1962.

Chao, Lin-shan. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

Dhrymes, Phoebus C. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

Hansen, Morris H. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

BIBLIOGRAPHY

Heady, P. F. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

Kabe, D. K. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

Kendall, Maurice G. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

Scoble, R. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

Kish, R. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

Kremer, J. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

Kotler, J. Statistical Theory and Applications. New York: McGraw-Hill, 1962.

## BIBLIOGRAPHY

## BOOKS

- Ackoff, Russell. Scientific Method: Optimizing Applied Research Decisions. New York: John Wiley & Sons, 1962.
- Anderson, T.W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley & Sons, 1958.
- Bliss, C.I. Statistics in Biology. New York: McGraw-Hill, 1970.
- Cacoullos, T. Discriminant Analysis and Applications. New York: Academic Press, 1973.
- Chao, Lincoln L. Statistics: Methods and Analyses. New York: McGraw-Hill, 1969.
- Dhrymes, Phoebus J. Econometrics: Statistical Foundations and Applications. New York: Harper & Row, 1970.
- Hansen, Morris H. et. al. Sample Survey Methods and Theory. New York: John Wiley & Sons, 1953.
- Heady, Earl O. Economic Models and Quantitative Methods for Decisions and Planning in Agriculture. Ames, Iowa: The Iowa State University Press, 1971.
- Kabe, D.G. and R.P. Gupta. Multivariate Statistical Inference. London: North-Holland Publishers, 1973.
- Kendall, Maurice G. Advanced Theory of Statistics. London: Charles Griffin & Company, 1946.
- Kendall, Maurice G. A Course in Multivariate Analysis. London: Charles Griffin & Company, 1965.
- Kish, Leslie. Survey Sampling. New York: John Wiley & Sons, 1965.
- Kmenta, Jan. Elements of Econometrics. New York: The Macmillan Company, 1971.
- Kotler, Phillip. Marketing Decision Making: A Model Building Approach. New York: Holt, Rinehart, and Winston, 1971.
- Krishnaiah, P.R. Multivariate Analysis II. New York: Academic Press, 1969.

- Mendenhall, William. An Introduction to Probability and Statistics. Belmont, California: Wadsworth Publishing Company, 1967.
- Overall & Klett. Applied Mathematical Analysis. New York: McGraw-Hill, 1972.
- Press, James. Applied Multivariate Analysis. New York: Holt, Rinehart, and Winston, 1972.
- Rao, Radhaskrishna C. Linear Statistical Inference and Its Applications. New York: John Wiley & Sons, 1965.
- Sills, David L. International Encyclopedia of Social Sciences. New York: The Macmillan Free Press, 1968.
- Tanur, Judith M., Frederick Mosteller, et. al. Statistics: A Guide to the Unknown. New York: Holden-Day, Inc., 1973.
- Wilks, Samuel. Mathematical Statistics. New York: John Wiley & Sons, 1962.

## ARTICLES

- Anderson, J.A. "Constrained Discrimination between K Populations," Royal Statistical Society, XXXI, Series B, 1969, 123-39.
- Araji A.A. and M. Finley, "Managerial Socioeconomic Characteristics and Size of Operation in Beef Cattle Feeding: An Application of Discriminant Analysis," American Journal of Agricultural Economics, LIII, 4, 1971, 648.
- Lurie, Irene. "Estimates of Tax Rates in the AFDC Program," National Tax Journal, XXVII, 7, 1974, 93-106.
- Morrison, Donald G. "On the Interpretation of Discriminant Analysis," Journal of Marketing Research, V6, 1969, 156.
- Prescott, James R. and William C. Lewis. "State and Municipal Locational Incentives: A Discriminant Analysis," National Tax Journal, XXII, 3, 1969, 402.

## PAMPHLETS

"South Dakota Indian Recipients on Social Welfare," State Department of Social Services (South Dakota), December, 1973.

"South Dakota Indian Recipients of Social Welfare," South Dakota State Department of Public Welfare, October, 1972.

"Research and Statistics Report," South Dakota Department of Public Welfare, October and December, 1972.

"Program Analysis Report," South Dakota Department of Social Services, September and December, 1973, March, April, and May, 1974.

"ADC Characteristics Study," South Dakota Department of Social Services, January, 1973.

APPENDIX



## APPENDIX I

1. Is your present income higher or lower than your past highest income?
  2. If your present income is higher (lower), how long has it been higher? \_\_\_\_\_ Months.
  3. Do you expect an increase (decrease) in your income in the next year? Yes No
  4. How strongly do you feel that your income will increase (decrease)?  

|                |   |   |   |   |   |                |
|----------------|---|---|---|---|---|----------------|
| 1              | 2 | 3 | 4 | 5 | 6 | 7              |
| Strong<br>Neg. |   |   |   |   |   | Strong<br>Pos. |
  5. How much do you think your income will increase (decrease)?
  6. Are you personally satisfied? Yes No
- APPENDIX**
7. What type of business do you do?
  8. What is your hourly wage? \_\_\_\_\_ \$/hr
  9. On the average, how many hours a week do you work? \_\_\_\_\_
  10. Is your present job generally the same type of employment that you had had in the past? Yes No
  11. Do you expect to stay in your present job for the next year? Yes No
- How strongly do you feel that you will stay in your present job for the next year?
- |                |   |   |   |   |   |                |
|----------------|---|---|---|---|---|----------------|
| 1              | 2 | 3 | 4 | 5 | 6 | 7              |
| Strong<br>Neg. |   |   |   |   |   | Strong<br>Pos. |

## APPENDIX I

1. Is your present income higher or lower than your past highest income?
2. If your present income is higher (lower), how long has it been higher? \_\_\_\_\_ Months.
3. Do you expect an increase (decrease) in your income in the next year? Yes No
4. How strongly do you feel that your income will increase (decrease)?
- |                |   |   |   |   |   |                |
|----------------|---|---|---|---|---|----------------|
| 1              | 2 | 3 | 4 | 5 | 6 | 7              |
| Strong<br>Neg. |   |   |   |   |   | Strong<br>Pos. |
5. How much do you think your income will increase (decrease)?
6. Are you presently employed? Yes No
7. What type of work do you do?
8. What is your hourly wage? \_\_\_\_\_ \$/hr
9. On the average, how many hours a week do you work? \_\_\_\_\_  
in the \_\_\_\_\_ hr/week
10. Is your present work generally the same type of employment that you have had in the past? Yes No
11. Do you expect to stay in your present job for the next year? Yes No
- |               |                |   |   |   |   |   |                |
|---------------|----------------|---|---|---|---|---|----------------|
| How strongly? | 1              | 2 | 3 | 4 | 5 | 6 | 7              |
|               | Strong<br>Neg. |   |   |   |   |   | Strong<br>Pos. |

12. Are you earning more in your present job than in your last steady job? Yes No How much more \_\_\_\_\_ \$/hr.
13. Do you expect an increase (decrease) in your wages this year? Yes No
14. If you are unemployed, how long has it been since you were steadily employed? \_\_\_\_\_ Months.
15. What type of work did you most often do?
16. What was your hourly wage in your last steady job? \_\_\_\_\_ \$/hr.
17. Do you think you will be re-employed this year? Yes No  
How strongly? 1 2 3 4 5 6 7  
Strong Neg. Strong Pos.
18. Do you think you will be re-employed at a higher (lower) wage than your last steady job? How strongly?  
Strong neg. 1 2 3 4 5 6 7 Strong pos.
19. Have you attempted to receive any employment counseling in the last year? Yes No
20. Have you attempted to receive any form of job training in the last year? Yes No
21. How long have you been receiving ADC payments?  
\_\_\_\_\_ Months.
22. Do you expect to continue (discontinue) receiving these payments for at least the next year? Yes No. How strongly?  
Strong Neg. 1 2 3 4 5 6 7 Strong Pos.
23. Do you currently have any savings? Yes No

24. If you have, or have had savings in the last year - have you had to spend any of these savings to meet current obligations? Yes No

25. Do you expect to save any of your income this year?  
Yes No How strongly?

Strong Neg. 1 2 3 4 5 6 7 Strong Pos.

26. Have you attempted to gain credit this year? Yes No

27. What type of lending institutions have you attempted to get credit from? Bank Loan Co. Credit Card Retail Store

28. Have you been turned down at any of these places in the last year? Yes No

29. Will you try (will not) to get additional credit in the next year? Yes No How strongly?

Strong Neg. 1 2 3 4 5 6 7 Strong Pos.

30. If you intend to gain additional credit, do you think that you will be able to obtain the credit you desire?

Yes No How strongly?

Strong Neg. 1 2 3 4 5 6 7 Strong Pos.

31. Are you married, single, divorced, separated, or widowed?

32. Do you expect a change in marital status in the next year? Yes No How strongly?

Strong Neg. 1 2 3 4 5 6 7 Strong Pos.

33. How many dependents are currently residing with you?

34. What are the ages of your dependents?

35. How many years of schooling have you completed?



Table 3-6

Pooled Dispersion Matrix  
(Symmetric Matrix)

---



---

|         |         |         |         |          |         |         |          |         |          |
|---------|---------|---------|---------|----------|---------|---------|----------|---------|----------|
| 2.57260 | -.66230 | 1.88338 | 2.73843 | .20700   | .59078  | 1.22977 | 1.84713  | .54925  | .03253   |
|         | 7.25530 | -.20111 | -.20240 | -.64811  | 1.66120 | .00553  | -1.13106 | 1.45806 | .29917   |
|         |         | 5.19896 | 3.60885 | -1.78544 | .27558  | 1.73189 | 1.90289  | .90151  | .40922   |
|         |         |         | 7.37917 | -.33512  | .63041  | 1.75576 | 3.20848  | 1.09548 | .72664   |
|         |         |         |         | 9.09346  | 1.72756 | -.47281 | -.41438  | .21548  | -1.27908 |
|         |         |         |         |          | 5.20479 | .86212  | 1.23576  | 3.55226 | .11668   |
|         |         |         |         |          |         | 5.16055 | 2.23889  | 1.32581 | .41392   |
|         |         |         |         |          |         |         | 7.29555  | 2.54172 | 1.19217  |
|         |         |         |         |          |         |         |          | 5.60430 | .85613   |
|         |         |         |         |          |         |         |          |         | 3.82691  |

---