THE RELATIVE EFFICIENCY OF SEVERAL PAROLE PREDICTION STRATEGIES.

JOHN M. SYROTUIK
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/4226
NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

Si il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ MICROFILMÉE TELLE QUE NOUS L'AVONS RÉCU
THE RELATIVE EFFICIENCY OF SEVERAL
PAROLE PREDICTION STRATEGIES

(John M. Syrotuik

A Dissertation
Submitted to the Faculty of Graduate Studies
Through the Department of Psychology
in Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
at the
University of Windsor

Windsor, Ontario
1977
ABSTRACT

The present investigation involved an evaluation of the relative efficiency of several multiple regression models and a sequential Bayesian technique as applied to the problem of predicting parole outcome. Both methods were applied several times with corresponding variations of relevant prediction parameters.

Each analysis was performed on the total sample of 205 parolees and was repeated four additional times on randomly selected subgroups. Each sub-group was composed of one quarter of the original subject pool and parole predictions were based on models derived using the remaining three-quarters of the total sample. Adaptation of the Tukey (1969) 'Jackknife' procedure as described here permitted an evaluation of the relative degree of shrinkage expected on the validation of each method.

The results indicated that there were no significant differences in predictive efficiency (on validation) between any of the strategies investigated. The regression and Bayesian techniques did exhibit different shrinkage patterns however, with the latter being less prone to overfitting on construction. The results were discussed in terms of their implications for further research in the area and parole decision policy.
ACKNOWLEDGEMENTS

I would like to thank my dissertation chairman, Dr. M. Morf, and the committee members Drs. J. Ager, F. Auld, V. Signorile and M. Starr for their assistance in the completion of the dissertation. I also wish to thank those individuals at the Windsor and London branches of the National Parole Service for their co-operation in obtaining the data.

Finally I express my sincere gratitude and appreciation to Rasa Jasevicius who contributed greatly to my understanding of the broader aspects of this dissertation.
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TABLE OF CONTENTS (cont'd)</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>b)</td>
<td>Sequential Bayesian Analysis</td>
<td>24</td>
</tr>
<tr>
<td>iii)</td>
<td>Moderator Variable Analysis</td>
<td>32</td>
</tr>
<tr>
<td>iv)</td>
<td>Summary of Analyses and Evaluation of Their Relative Efficiency</td>
<td>37</td>
</tr>
<tr>
<td>2.</td>
<td>Methodological Points of Importance</td>
<td>39</td>
</tr>
<tr>
<td>i)</td>
<td>The Tukey 'Jacknife' Procedure</td>
<td>39</td>
</tr>
<tr>
<td>ii)</td>
<td>Collection and Preparation of Data Base</td>
<td>40</td>
</tr>
<tr>
<td>iii)</td>
<td>Operational Definition of Criterion and Moderator Variable</td>
<td>41</td>
</tr>
<tr>
<td>IV</td>
<td>RESULTS</td>
<td>43</td>
</tr>
<tr>
<td>1.</td>
<td>Preliminary Analyses of the Data Base</td>
<td>43</td>
</tr>
<tr>
<td>2.</td>
<td>Evaluating the Relative Efficiency of Prediction</td>
<td>44</td>
</tr>
<tr>
<td>V</td>
<td>DISCUSSION</td>
<td>60</td>
</tr>
<tr>
<td>1.</td>
<td>Conclusions Regarding the Relative Efficiency of Prediction</td>
<td>60</td>
</tr>
<tr>
<td>2.</td>
<td>Implications of Results for Parole Decision Policy</td>
<td>64</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>REFERENCES</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>VITA</td>
<td>103</td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Criterion Correlations, MCR's, and Rank Order (by Technique) of Final Nineteen Item Predictor Pool</td>
<td>45</td>
</tr>
<tr>
<td>2</td>
<td>Summary Descriptions of Final Nineteen Item Predictor Pool</td>
<td>46</td>
</tr>
<tr>
<td>3</td>
<td>Multiple Regression Equations Generated on the Basis of the Entire Sample</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>MCR and Multiple Correlation Coefficients of Regression Analyses Computed on the Basis of the Entire Sample</td>
<td>49</td>
</tr>
<tr>
<td>5</td>
<td>Average MCR and Multiple Correlation Coefficients of Regression Analyses Computed on the Basis of the Four Sample Pieces</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>Upper Limits (p &gt; 95%) of Multiple Correlation Coefficients of Regression Analyses Estimated on the Basis of the Four Sample Pieces</td>
<td>52</td>
</tr>
<tr>
<td>7</td>
<td>MCR and Multiple Correlation Coefficients of Bayesian Analyses Computed on the Basis of the Entire Sample</td>
<td>53</td>
</tr>
<tr>
<td>8</td>
<td>Average MCR and Multiple Correlation Coefficients of Bayesian Analyses Computed on the Basis of the Four Sample Pieces</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>Upper Limits (p &gt; 95%) of Multiple Correlation Coefficients of Bayesian Analyses Estimated on the Basis of the Four Sample Pieces</td>
<td>56</td>
</tr>
<tr>
<td>10</td>
<td>Average Multiple Correlation Coefficients of Moderator Variables Analyses Computed on the Basis of Two Sets of Four Sub-group Pieces</td>
<td>58</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1. Purpose of the Present Study

The primary goal of the present study was to determine the relative efficiency of several strategies of predicting parole outcome. The two general procedures employed were multiple linear regression and Bayesian analysis. Specifically, four different multiple regression models and two Bayesian methods were investigated. The four regression models were the general model, the forward model, the stepwise model and a unit weighting model. The most efficient regression model was also compared to the most efficient Bayesian method in an exploratory analysis utilizing a moderator variable.

Research related to parole outcome prediction has a long history—dating back to 1923—and it was necessary to consider this body of research in order to clearly develop and present the rationale of the present investigation. Before doing so, however, an attempt will be made to clarify certain aspects of the parole system and more fully describe the issue being considered in this study.

2. Parole: A Definition

Parole is a type of conditional release from prison whereby an offender serves part of his sentence in the community. In Canada, the decision as to whether or not an individual is granted parole is made by the National Parole Board which was officially established in 1959 by the Parole Act (see Appendix A). While on parole an
individual must abide by certain restrictions and conditions (see Parole Agreement, Appendix B) which include varying degrees of supervision by a Parole Officer. An offender may lose his parole and be returned to prison in one of two ways. First, parole may be suspended and then revoked for violation of the parole agreement. (A parolee may also have his parole revoked if he commits an offence while on parole.) Secondly, parole may be lost by forfeiture which occurs automatically if a parolee is convicted of an indictable offence punishable by a term of imprisonment for two years or more.

What was predicted in the present study, therefore, was whether or not an individual completes the parole agreement successfully. In the last fifty years numerous approaches—varying greatly in complexity and sophistication—to this prediction problem have been investigated. The most important of these procedures will now be considered.
CHAPTER II
REVIEW OF PAROLE PREDICTION METHODOLOGY

1. Predictive Efficiency and Base Rates

Grygier, Blum and Porebski (1971) point out that the typical sample of Canadian parolees has about an 80% success rate; that is, about 80% of all individuals on parole complete their paroles without forfeiture or revocation. With this base rate information alone, one could predict that all parolees in a sample will complete their paroles successfully and be correct approximately 80% of the time. It is apparent that any prediction technique will be worthwhile only to the extent that it improves upon the accuracy of predictions made using the base rate alone.

As has been pointed out by numerous investigators (e.g., Blumstein & Larson, 1971) predictions of this nature involve two types of potential errors, namely: (i) predicting parole violation where it does not occur (false positives) and, (ii) predicting non-violation where parole violation occurs (false negatives). If parole outcome predictions are based on the base rate alone, the number of false positives equals zero but the false negative rate is maximized. Furthermore, the closer the base rate is to .50, the greater the number of false negatives.

It follows from the above that the utility of any method of parole prediction may be viewed as the extent to which it maintains the zero false positive rate while decreasing the number of false negatives associated with base rate predictions. In theory, the
degree to which any particular method of prediction accomplishes these ends will tend to increase as the base rate approaches .50 (Ohlin & Lawrence, 1952; Simon, 1971). This is expected in light of the corresponding increase in the variance of the criterion and the greater likelihood of finding relevant predictors. As Simon (1971) points out, however:

"Although theoretically it would seem that a base rate of .5 would be the ideal, examination of many studies with base rates (for failure) of between about .25 and .50 suggests that within this range at any rate, the base rate is a very minor factor in determining the power that is actually obtained in practice". (p. 26)

All of the prediction methods to be discussed below involve assigning each parolee to one of a number of risk categories characterized by different base rates. The extent to which the base rates of these categories differ from the base rates of the sample as a whole will determine the utility of the prediction method. Although this view of predictive efficiency might appear to differ from the above formulation involving the false positives and negatives of base rate predictions, it is in fact equivalent. This is illustrated by the extreme and simplified case where individuals are assigned to one of two categories with 0% and 100% success base rates. Predictions based on category membership here would maintain the zero false positives of base rate predictions while reducing to zero the number of false negatives. Categories with less extreme base rates would accomplish this to a lesser extent.
The prediction system discussed here represents the simplest two-category case and assigning category membership in this case is operationally equivalent to predictions of success and failure on parole. The rationale of the above discussion, however, applies with equal validity to predictive systems where several risk categories are involved. In fact, the generation of several predictive categories with widely divergent base rates can be seen as the primary goal of underlying applications of the methods which will now be considered.

The main purpose of the section which follows will be to describe the most important parole prediction methods which have been employed to date. Although these descriptions will be accompanied by general evaluative comments, the relative efficiency of these methods will be dealt with in greater detail at a later point.

2. Linear Methods of Prediction

(i) Unweighted Points Score Analysis

In 1928 Burgess, based on the earlier work of Hart (1923) and Warner (1923), developed a simple points score system of parole prediction. This system involved assigning a 1 or 0 for each of 21 items (e.g., first offender vs. repeat offender) and totalling these scores for each parolee. The total score was then used to assign the individual to one of nine risk categories. These categories which represented different portions of the 0 to 21 range had violation rates ranging from 1.5% (scores 15-21) to 76.0% (scores 2-4).
The Burgess approach has been evaluated in numerous studies (e.g., Tibbitts, 1931, 1932; Vold, 1931a, 1931b) which on the whole have indicated that the method is reasonably reliable and accurate enough to be developed for application in the correctional field. At the same time, however, these studies indicated that the accuracy of tables based solely on background factors might develop severe limitations over time under certain circumstances (e.g., changing administrative policies). Furthermore, tables of this type appeared to have limited generalizability across different populations of parolees.

(ii) Weighted Points Score Analysis

Another method of major importance in early parole prediction research was that developed by the Glueck research team (e.g., Glueck, S & E, 1929; Glueck, S., 1932). This approach differs from the Burgess technique in that fewer items are utilized (6 as opposed to 21) and these items are selected on the basis of their association with the criterion. As in the Burgess method each individual is assigned a score for each item and these are totalled for purposes of assigning him to a particular risk category. Unlike the Burgess method, however, the weights of each item are not 0 or 1 but rather the % of total failures found in each subcategory of the item. To illustrate, consider the example item, first offender versus repeat offender. If an individual is a first offender and if 20% of all parole violators in the sample are first offenders, then the individual receives a score of .20 for that item. In this
case if the individual was a repeat offender he would receive a score of .80 for the item. The Glueck advocated their approach on the grounds that it represented an economy of both time and money—since it used fewer items—with minimal loss in predictive accuracy.

The Glueck method has been applied extensively in attempts to predict juvenile delinquency and has been the subject of numerous criticisms (see Simon, 1971, p. 37). With respect to parole prediction, however, applications of the method (e.g., Hakeem, 1945; Schwartz, 1949) have yielded results which are comparable to the Burgess technique. As Grygier (1956) points out, however, there is no logical reason why the percentage of offenders who possess an attribute should be the weight of that attribute in a prediction battery.

(iii) Multiple Linear Regression

Although regression might be viewed as a weighted points score analysis, it will be considered as a completely separate procedure in the light of the rather complex manner in which item weights are derived. At this point only the most important applications of this technique will be considered, while a more detailed discussion of the method is presented in Appendix D.

Kirby (1954) initially suggested that multiple regression be applied to the prediction of parole outcome. He recommended this approach on the ground that items were uniquely weighted in the light of their relationships to other predictors as well as to the
criterion. Mannheim and Wilkins (1955), Wilkins (1955) were the first investigators to conduct a thorough study of the regression procedure as applied in the field of corrections. The predictive strategy which they developed has been extensively applied in the parole system by the California Department of Corrections.

The research of the latter organization lead to the development of the well-known Base Expectancy Score (BES) system. This 12-item prediction system is essentially the same as the Burgess and Glueck procedures except for the fact that it utilizes the beta weights derived from multiple regression.

Among the more interesting studies employing multiple regression is an investigation conducted by Gough, Wenk and Rozynko (1965). These authors developed and cross validated six multiple regression equations which utilized various combinations of (i) the California Personality Inventory (CPI), (ii) the Minnesota Multiphasic Personality Inventory (MMPI), and (iii) the Base Expectancy Score (BES). Gough, et al. (1965) found that cross-validation of all equations yielded predictions more accurate than those which would be obtained from the base rates alone. The most accurate predictions were those generated by the equation which utilized both the CPI and the BES. The addition of the MMPI to this equation did not improve prediction.

3. Non-linear Methods of Prediction

All of the methods discussed to this point assume that it is appropriate to add weighted values of each predictor and that the score so generated is the best estimate of the criterion. Strict
adherence to such an assumption, however, preempts consideration of interactions among the independent variables which many investigators suggest would significantly improve predictive efficiency. In an attempt to capitalize on this potential source of information, several closely related non-linear hierarchical methods have been investigated; the most important of these are: (i) association analysis, (ii) predictive attribute analysis, (iii) configural analysis, and (iv) mean cost rating (MCR) analysis.

All of these techniques proceed in a step-by-step (hierarchical) fashion whereby the sample is successively divided into n subgroups each division being based on a single predictor. The predictors selected for this purpose have a high association with the criterion of parole outcome but, as will be discussed below, the procedures differ with respect to the manner in which this association is evaluated.

(i) Association Analysis

This method was originally developed by Williams and Lambert (1960) for studying communities of plants and was subsequently applied to the prediction of parole outcome by Wilkins and MacNaughton-Smith (1984). Using this method one can derive n subgroups in the hierarchical manner described above with each successive subdivision being based on the possession or lack of a predetermined attribute. The attribute which is chosen for this purpose is the one which is most closely associated with the other attributes describing the group. For each attribute, therefore, all the $\chi^2$'s for the $2 \times 2$
tables relating it to all other attributes are summed and the one having the largest $\chi^2$ is chosen for dividing the group.

It is apparent that this method is a descriptive rather than predictive one and as such one would not expect high predictive power. At the same time, however, if the attributes upon which divisions are made are themselves predictors, it is very likely that the classification will be predictive also. This method has been recently applied in this manner to the prediction of parole outcome (e.g., Chapman & Sinclair, 1973). The initial results indicate that while construction sample subgroups appear to have some predictive utility, there may be problems associated with the stability of the method across samples (Simon, 1971; Fildes & Gottfredson, 1972).

(ii) Predictive Attribute Analysis

Like association analysis, predictive attribute analysis (developed by MacNaughton-Smith, 1965) proceeds hierarchically by forming groups defined by the possession or lack of a specified attribute. The attribute chosen is the one most strongly associated with parole outcome of a particular level of the hierarchy.

In applying this method, one would proceed in the following manner. Chi square statistics are first computed to measure the association between parole outcome and any number of attributes. The initial sample is then divided on the basis of the attribute with the largest $\chi^2$ and all $\chi^2$ are then re-computed for each subgroup. These subgroups are subsequently divided on the basis of
the attribute with the largest $x^2$ (not necessarily the same ones) and this process is continued until some arbitrarily determined criteria of $x^2$ or subgroup size is reached. This procedure has been evaluated in a number of studies (e.g., Vichert & Zahnd, 1965; Grygier, 1966; Carney, 1967; Grygier, Blum & Forebski, 1971) and the method appears to be a viable one despite the danger of capitalizing on chance factors to which it is particularly vulnerable. (This risk is inherent in the process of selecting the best predictor at every stage of splitting.)

(iii) Configural Analysis

This approach was proposed by Glaser (1962) and is quite similar to MacNaughton-Smith's predictive attribute analysis. Glaser (1962) suggested that one should determine which of the input variables best divides the sample into subgroups with widely divergent failure rates. Following the division of the sample on the basis of this variable, the process is repeated several times in each subsequent subgroup until some arbitrary minimum group size is reached.

The primary difference, therefore, between this procedure and predictive attribute analysis is that the latter utilizes $x^2$ to test for the best sample splits while the configural approach uses the divergence in the failure rates of extreme variable categories. (Such divergence may be assessed in a different manner using MCR analysis but discussion of this procedure will be dealt with below as a separate technique). Another important difference between these two methods is that only attributes (i.e., dichotomous variables)
may be utilized in predictive attribute analysis while in configural analysis, variables with two or more discrete values can be employed. This apparent advantage of the configural approach, however, may render it more susceptible to sampling fluctuations (see Simon, 1971, p. 104).

In the light of the results obtained to date (e.g., Babst, Gottfredson & Ballard, 1968; Babst, Inciardi & Jamon, 1971; Inciardi, 1971; Inciardi & Babst, 1971) the configural approach to parole prediction would appear to be a worthwhile procedure. It should be noted here that this is one of the few techniques which has seen some application in practice. Parole prediction tables have been developed using configural analysis for Wisconsin (Babst & Kasado, 1964), the District of Columbia (Pownall & Karaicki, 1966), Massachusetts (Metzner & Weil, 1963), and Los Angeles County, California (Adams & Thompson, 1963).

When Glaser introduced configuration analysis in 1962, he described it as a less technical adaptation of Mean Cost Rating (MCR) analysis developed almost a decade earlier by Duncan, Ohlin, Reiss and Stanton (1953). Because the latter method relates in a significant way to sequential bayesian analysis—to be employed in the present study—it will be discussed in some detail at this point.
(iv) Mean Cost Rating Analysis

In mean cost rating (MCR) analysis, one proceeds by developing 'cost-utility curves' (C-U curves) for all items and splitting the sample on the basis of the 'best' curve. This process is continued in a hierarchical fashion with subdivisions being made on the basis of successive families of C-U curves produced from derived subgroups. At any level of the hierarchy, the C-U curve for a variable represents the functional relationship between 'cost' \( C_i \) and 'utility' \( U_i \) (to be defined below) over the possible ordinal values of that variable; that is, \( C_i = f(U_i) \).

To explain, suppose that a decision rule is adopted to 'accept' cases above score \( i \) and 'reject' cases below that score. In this instance 'cost' is defined as the proportion \( C_i \) of parole successes rejected at score \( i \) and 'utility' is the proportion \( U_i \) of failures rejected. Duncan, et al., (1953) point out that the best variables for subdivision of the sample at any point will be the one which has the lowest \( C_i \) values for all values of \( U_i \). Since the mean cost over utility interval \( U_{i-1} \) to \( U_i = (C_i + C_{i-1}) / 2 \), then the mean cost over all intervals of utility is given by the weighted mean or means of the individual intervals. (The weights here are the magnitudes of the utility intervals \( U_i - U_{i-1} \).) Hence,

\[
\text{Mean Cost (MC)} = \frac{1}{2} \sum_{i}^{k} (C_i + C_{i-1})(U_i - U_{i-1})
\]

In order to generate an index of selectivity on a 0 to 1 scale with positive increments reflecting increased selectivity, MCR is defined as,
MCR = 1 - 2 MC
= 1 - \sum_{i}^{k} \left( \frac{C_i}{i} + \frac{C_{i-1}}{i-1} \right) (U_i - U_{i-1})
= \sum_{i=1}^{k} \frac{C_i U_{i-1} - C_{i-1} U_i}{i}
\text{(see Inciardi, Babst & Koval, 1973)}

Duncan, et al. (1953) point out that if MCR is computed from just three values \((0,0), (U_j, C_j)\) at any given point, and \((1,1)\) it can be shown that \(MCR = U_j - C_j\). This represents the attribute case and would be numerically equivalent to Glaser's (1962) criteria of the difference between the failure rates of variable subgroups.

The major limitation of the MCR approach are (i) the variable categories must form a logical continuum with respect to the predicted variable (not relevant to the attribute case) (Inciardi, Babst & Koval, 1973) and (ii) MCR is independent of subgroup size and may remain large even though it refers to a subgroup too small to show a statistically significant difference (Simon, 1971). Despite these problems, however, MCR analyses have yielded results comparable to the other non-linear hierarchical methods (Simon, 1971; Inciardi, Babst & Koval, 1973).

The above discussion completes our consideration of the methods which have been employed to predict parole outcome. It would appear rather surprising that a Bayesian strategy has not as yet been investigated with respect to the parole prediction paradigm. The application of such a procedure in the present study, therefore, represents an extension of present knowledge in this field.

All of the techniques discussed above would appear to yield some improvement over base rate predictions. Before proceeding
to a detailed discussion of the Bayesian method, therefore, let us consider the relative predictive efficiency of these procedures which have already been applied.

4. Evaluating Predictive Efficiency

(i) The Relative Efficiency of Prediction Methods

The problem of measuring the predictive power of an equation or table for comparison purposes may be approached in several ways. For methods such as regression, where a continuous probability score is generated, the correlation between the predicted and observed score is the obvious approach. For techniques which generate tables (e.g., the non-linear hierarchical methods) the correlation coefficient $\phi = \frac{x^2}{n}$ for a $2 \times k$ table (see Simon, 1971; Appendix A).

Another measure which can be employed for evaluating the efficiency of prediction tables is the MCR discussed above. This index is influenced not only by the amount of divergence achieved between risk groups (as is $\phi$) but also by the degree to which they are in strict order of failure rates. $\phi$ has the disadvantage of being influenced by the ratio of successes to failure in the sample while MCR is independent of this factor. Both $\phi$ and MCR are also influenced by the number of risk groups derived.

Another measure of predictive efficiency is Ohlin and Duncan's (1949) 'Index of Predictive Efficiency' (IFE). This index is derived by dividing the % of error decrease using the prediction method by the % error decrease using the base rate. The IFE has been extensively utilized despite the fact that it has the serious

---

*The term efficiency is used here in the pragmatic as opposed to the statistical sense.
disadvantage (discussed by Simon, 1971, p.22) of generating a zero efficiency rating in the circumstances where risk group failure rates and the overall failure rate are both greater or less than .5.

At an earlier point in this paper, it was suggested that the utility of any prediction method could be viewed as the extent to which that technique maintained the zero false positive rate and decreased the number of false negatives associated with base rate predictions. It was proposed that the degree to which this was achieved was directly related to the extent to which the base rates of the derived risk groups were different from the base rates of the sample as a whole. As noted above, both the MCR and $\phi$ (and $r$) are influenced by this divergence and as such, represent measures of efficiency consistent with our earlier formulation.

The most important studies comparing methods of predicting parole outcome have been conducted by (i) Vold (1936), who compared the Glueck and Burgess techniques, (ii) Cohlin and Duncan (1949) who compared—among others—Glueck and Burgess' prediction tables, (iii) Ward (1968), who compared the Burgess, Glueck, multiple linear regression and discriminant function procedures, (iv) Mannheim and Wilkins (1955) who compared multiple linear regression and point scoring systems, (v) Gottfredson and Beverly (1962) who compared multiple regression, association and predictive attribute analysis, (vi) Babst, Gottfredson and Ballard (1968) who compared linear regression and configurational analysis, and (vii) Grygier (1969) who compared discriminant function, association, predictive attribute, and points score analyses. The most exhaustive investigation of
this type was conducted by Simon (1971, 1972) who compared linear regression, association, predictive attribute, configuration, mean cost rating and two points scoring analyses.

In all of the above studies, no significant differences were found with respect to the predictive power of the instruments compared. The results of the Simon (1971) investigation typify the findings of the other studies cited and will be dealt with in some detail at this point.

Simon (1971) conducted five multiple regression analyses, two association analyses, five predictive attribute analyses, two points score analyses, and a configuration and mean cost rating analysis on both construction and validation samples. The regression analyses generated multiple correlation coefficients on construction ranging from .37 to .55 which upon validation shrank to between .13 and .18. The two association analyses had rather low predictive power on construction (MCR = .24, \( \phi = .21 \); MCR = .25, \( \phi = .23 \)) which dropped somewhat on validation (MCR = .11, \( \phi = .14 \); MCR = .19, \( \phi = .21 \)). Three of the five predictive attribute analyses produced MCR's on construction of between .50 and .57 but these dropped to zero in the validation sample. The two other predictive attribute analyses, the configuration analysis and the MCR analysis, produced MCR's on construction ranging between .26 and .45 which dropped to between .17 and .31 on validation. As Simon (1971) points out, 'The conclusion to be drawn from this examination of statistical methods for combining data is that in practice all of them work
about equally well. The exception is association analysis, which is not designed to be predictive, though in some situations it may be.' (p. 154).

The evidence discussed above would suggest that there are important source(s) of error which act to limit the predictive efficiency of all procedures compared. At this point, therefore, let us consider the most important of these (error) factors as they pertain to the present study.

It should be noted here again, that the purpose of the present investigation was to compare the efficiency of a number of different predictive strategies and not to develop a maximally efficient procedure. Because all methods, in application, were potentially subject to the same sources of error, the latter does not in any way affect the appropriateness of these comparisons.

(iii) Factors Limiting Predictive Efficiency

There are several important, and unfortunately, unavoidable factors which limited the degree of predictive accuracy obtained in the present investigation. One major source of error was to be found in the data itself. A number of investigators (e.g., Huff, 1936; Vold, 1949; Simon, 1971) have questioned the reliability and validity of predictor variables recorded from the parole board files of offenders. In the same way that reliability represents the upper limit of a variable's validity, it likewise places a ceiling on the efficiency of a predictive instrument.

The measurement of the criterion variable (i.e., parole outcome) represents another source of error. First of all, it is apparent
that measures of the criterion were not perfectly valid. Specifically, many crimes or other behaviours warranting loss of parole often go undetected and there can be no doubt that parole was sometimes revoked or forfeited where such action was not justified.

Certain questions can also be raised regarding the reliability of the criterion. Variation among parole officers with respect to their interpretation and application of parole policy constitutes an additional source of criterion measurement error.

In almost all studies of parole prediction, including the present one, a dichotomous criterion of success-failure was utilized. Simon (1971) has suggested that such a dichotomy may represent an over-simplification of criminality which is more appropriately viewed in terms of a continuous scale. The extent to which information is lost by the bipolar definition also reduces predictive accuracy. It should be pointed out here, however, that the problem has traditionally been approached in this manner because it is not apparent how a more useful continuous scale would be defined and none has, in fact, been forthcoming.

Another source of prediction error was that associated with the fluctuation of statistical parameters across construction and validation samples. The efficiency of a prediction system will undergo shrinkage when applied to the validation sample as a function of the magnitude of such fluctuations which are due to sampling error. Tukey's (1969) 'jackknife' procedures—to be described in later section—was employed in the present study in an attempt to reduce such error.
The most serious source of error, however, was that associated with the passage of time. Various forces (e.g., political, economic, etc.) were expected to alter the relationship between predictors and criterion over time—possibly even affecting a redefinition of the criterion. To date, no prediction strategy has been able to significantly reduce time errors—although a number have been proposed, e.g., see Fano’s (1961) discussion of ensemble averages—and they inevitably affected predictive accuracy in the present study.

A final limitation to the efficiency of any predictive strategy is an absolute one and is mentioned here in the interest of a complete discussion. This limitation represents the degree to which quantified variables can be expected, even under ideal conditions, to predict the criterion in the population. In other words, even with perfectly reliable and valid data, and assuming no time, sampling or similar errors, there may still be inherent limitations in a quantified prediction system.

Bearing the above limitations in mind, let us now consider the prediction methods employed in the present study.
CHAPTER III

METHOD

1. Prediction Strategies to be Investigated

(i) Multiple Regression Analysis

As discussed at an earlier point, multiple regression has, in fact, been applied to the problem of predicting parole outcome (e.g., Mannheim & Wilkins, 1955; the California Base Expectancy Tables). In the present study, a more thorough investigation of the general, forward and stepwise regression models was undertaken with variations (to be discussed below) being made with respect to levels of significance and the number of predictors. The question can be raised here, however, as to whether or not the optimal weighting system derived using any regression model is worthwhile. Our earlier discussion of studies comparing points score and regression techniques, would in fact, suggest that a unit weighting system may be just as efficient and much more simply derived. This speculation is supported by several other studies (e.g., Wesman & Bennett, 1959) which indicate that one can depart from the optimal weights of a least squares solution to a unit weighting system with minimal loss in efficiency as a function of (i) the ratio of observations to variables, (ii) the presence or absence of suppressor variables (Schmidt, 1971), and (iii) the intercorrelations among variables (Lawshe & Schäcker, 1951). These three factors all affect the magnitude of the standard error of the derived beta weights and it is clear that use of unit weights

21
is most appropriate where such error is large. In the present study, where there are numerous sources of error (as described), a unit weighting system may prove to be as efficient as the much more complex regression procedures.

(ii) Bayesian Analysis

(a) Non-Sequential Bayesian Analysis

For purposes of outlining the manner in which a non-sequential Bayesian method can be applied to the problem of predicting parole outcome, the following definitions are provided:

\[ p(S) = \text{the base rate probability of success on parole.} \]

\[ p(F) = \text{the base rate probability of failure on parole.} \]

\[ p(A|S) = \text{the conditional probability of attribute A given success on parole.} \]

\[ p(A|F) = \text{the conditional probability of attribute A given failure on parole.} \]

\[ p(S|A) = \text{the conditional probability of success on parole given attribute A.} \]

\[ p(F|A) = \text{the conditional probability of failure on parole given attribute A.} \]

\[ p(S \wedge A) = \text{the joint probability of success on parole and attribute A.} \]

\[ p(A) = \text{the probability of attribute A.} \]

Given that attribute A is related to parole outcome S, the rule of conditional probability allows one to modify the base rate probability \( p(S) \) in the light of the information provided by this relationship; that is,

\[ p(S|A) = \frac{p(S \wedge A)}{p(A)} \]
Using the rule of multiplication (i.e., \( p(S \cap A) = p(S) \cdot p(A|S) \)) and the rule of elimination (i.e., \( p(A) = p(S) \cdot p(A|S) + p(F) \cdot p(A|F) \)) this function may be expressed as,

\[
p(S|A) = \frac{p(S) \cdot p(A|S)}{p(S) \cdot p(A|S) + p(F) \cdot p(A|F)}
\]

where \( S \) (success) and \( F \) (failure) are mutually exclusive events.

This formula is known as Bayes theorem and as shown, is a trivial consequence of the rule of conditional probability. It is apparent, however, that it can be employed as a method of predicting parole outcome. Specifically, one can utilize Bayes theorem to generate the a posteriori probabilities \( p(S|A) \), \( p(F|A) \) given a prior or base rate probabilities \( p(S) \), \( p(F) \) and the conditional probabilities \( p(A|S) \), \( p(A|F) \). The corresponding formula for generating the a posteriori probability of failure would be,

\[
p(F|A) = \frac{p(F) \cdot p(A|F)}{p(F) \cdot p(A|F) + p(S) \cdot p(A|S)} = 1 - p(S|A)
\]

Both of the above formulae apply to one pole of the attribute \( A \). If in fact the opposite pole of this attribute is relevant then the corresponding formulae are,

\[
p(S|\overline{A}) = \frac{p(S) \cdot (1 - p(A|S))}{p(S) \cdot (1 - p(A|S)) + p(F) \cdot (1 - p(A|F))}
\]

and

\[
p(F|\overline{A}) = \frac{p(F) \cdot (1 - p(S|F))}{p(F) \cdot (1 - p(A|F)) + p(S) \cdot (1 - p(A|S))}
\]

Utilization of any of the above formulae for predicting parole outcome would constitute a non-sequential Bayesian approach. (It is also noted here that, rather than a single attribute, \( A \) could
in fact represent a vector or pattern of attributes. The method would still be non-sequential, however.

(b) Sequential Bayesian Analysis

The method described can easily be extended so that it becomes a sequential method in which the a posteriori probabilities \( p(S|A) \), \( p(F|A) \) are generated in a step-by-step fashion on the basis of not one attribute but several attributes \( a_i \) in an attribute vector \( A \). In such an analysis a critical value (e.g., .90) is set and if this value is exceeded at any point by the a posteriori probabilities, the procedure is terminated. The sequential method, therefore, will not necessarily use all of the information available in attribute vector \( A \) whereas, the non-sequential method will always do so.

An expanded version of Bayes formula can be used to represent the sequential technique (Lasker, 1971) and for the case of parole success it takes the form,

\[
p(S|a_1, a_2 \ldots a_k) = \frac{p(S) \cdot p(a_1|S) \cdot p(a_2|S) \ldots p(a_k|S)}{p(S) \cdot p(a_1|S) \cdot p(a_2|S) \ldots p(a_k|S) + p(F) \cdot p(a_1|F) \cdot p(a_2|F) \ldots p(a_k|F)}
\]

Similar formulae, corresponding to those discussed earlier in reference to sequential analysis, can be generated here for the a posteriori probability of failure as well as for those cases where the opposite attribute poles are relevant.

In order to illustrate the manner in which the above formula would be employed for purposes of predicting parole outcome consider the following numerical example. Suppose that initially the following probability estimates are obtained from a sample:
\[ p(S) = .70, \, p(F) = .30, \, p(a_1 | S) = .95, \, p(a_2 | F) = .55, \, p(a_1 | S) = .85, \, p(a_2 | F) = .55. \]

If an individual possesses both attributes \( a_1 \) and \( a_2 \) the computations would involve two steps. The first step involves modifying the base rate probability estimates on the basis of attribute \( a_1 \). (Note: The method of selecting the order of the attributes considered is rather complex and will be dealt with below) in the following manner:

\[
p(S | a_1) = \frac{(.70)(.95)}{(.70)(.95) + (.30)(.55)} = .80
\]

The probability of success on parole for that individual, given attribute \( a_1 \) is therefore .80. Proceeding to consider attribute \( a_2 \) in step two one computed the probability of success as:

\[
p(S | a_1, a_2) = \frac{(.80)(.85)}{(.80)(.85) + (.20)(.55)} = .86
\]

The latter probability estimate can also be obtained directly using the Lasker (1971) formula as follows:

\[
p(S | a_1, a_2) = \frac{(.70)(.95)(.85)}{(.70)(.95)(.85) + (.30)(.55)(.55)} = .86
\]

As was the case with the other methods of parole prediction discussed earlier, the Bayesian technique can be evaluated in terms of the extent to which it places individuals in subgroups characterized by base rates which are maximally different from the base rates of the sample as a whole. In terms of the above discussion, it is clear that more efficient prediction will be obtained when the a posteriori probabilities of success and
failure are maximally different from the corresponding a priori probabilities. Let us now consider the manner in which one selects predictor attributes which will generate widely diverging a posteriori probabilities.

Slovic and Lichtenstein (1971) point out that Bayes theorem may be re-written as (using present notation),

\[
\frac{p(S|a_1)}{p(F|a_1)} = \frac{p(a_1|S)}{p(a_1|F)} \cdot \frac{p(S)}{p(F)}
\]

OR

\[
A = LR \cdot A'
\]

where:  
\(A\) = the posterior odds  
\(LR\) = the likelihood ratio  
\(A'\) = the prior odds

These authors propose that the degree to which the prior odds change depends on the likelihood ratio which represents an index of data diagnosticity or importance analogous to the weights employed in regression models.

In the present study, where two mutually exclusive hypotheses are being evaluated on the basis of their functional relationship to various attributes, the likelihood ratio may not be the most appropriate index. It is proposed here that the most important and straightforward measure of prediction importance is the magnitude of the absolute difference \(|p(a_1|S) - p(a_1|F)|\).

To illustrate consider an example where \(p(a_1|S) = .9\), \(p(a_1|F) = .8\) and \(p(a_2|S) = .3\), \(p(a_2|F) = .2\) for attributes \(a_1\) and \(a_2\). Using the likelihood ratios alone, \(a_1\) (LR = 1.1250) appears to be a poorer
predictor than \( a_2 \) (LR = 1.5000). This obscures the fact that only one pole of the attribute is being considered, however. If the likelihood ratios of each pole are weighted in terms of the proportion of subjects to which they apply the statistics so generated indicate that \( a_1 \) is a better predictor than \( a_2 \) (1.0125 versus .4500) while for the opposite poles \( \bar{a}_2 \) is more important than \( \bar{a}_1 \) (.6125 versus .0500). Of greater interest here is the fact that the sum of the weighted likelihood ratios of both poles of attributes \( a_1 \) and \( a_2 \) indicates that they are equivalent predictors; that is,
\[
LR_{a_1}^w + LR_{\bar{a}_1}^w = LR_{a_2}^w + LR_{\bar{a}_2}^w = 1.0525
\]

where: \( LR_{a_1}^w \) = the proportionately weighted likelihood ratio of the positive pole of attribute \( a_1 \).

The information needed to evaluate the relative importance of predictors \( a_1 \) and \( a_2 \) can be obtained in a much simpler fashion, however, by simply computing the absolute differences \([p(a_1|S) - p(a_1|F)] \) and \([p(a_2|S) - p(a_2|F)] \). In the case of the above example, the quantities are both equal to .1 indicating that the predictors are of equal importance. Similar information will be obtained if one considers the opposite poles of the attribute; that is, the differences \([p(\bar{a}_1|S) - p(\bar{a}_1|F)] \) and \([p(\bar{a}_2|S) - p(\bar{a}_2|F)] \) are likewise equal to .1.

It will be recalled that the MCR of a predictor was earlier defined in the attribute case to be MCR = \( U_j - C_j \) where \( U_j \) is the proportion of failures rejected at score \( j \) and \( C_j \) is the proportion
of successes rejected at that score. If all cases characterized by attribute $a_j$ are rejected it is clear that $U_j = p(a_j|F)$ and $C_j = p(a_j|S)$ and, as such, $|p(a_j|F) - p(a_j|S)| = U_j - C_j = MCR$. The difference between the endorsement proportions of attribute $a_j$ therefore is equal to the MCR of $a_j$ and provides the same information regarding predictor importance as the proportionally weighted sum of the likelihood ratios of each pole of $a_j$.

In addition to the manner in which predictors are selected, MCR analysis is related to the sequential Bayesian approach in a number of other important respects. Earlier it was noted that in MCR analysis the sample is split in a hierarchical manner on the basis of predictors which have the highest MCR in successive subgroups. In terms of the above discussion of attributes, the first sample split would be made on the basis of the absolute difference

$|p(a_j|S) - p(a_j|F)|$ where this difference is greatest for the attribute $a_j$. Because this quantity is computed on the basis of the entire sample, by definition it follows that the base rates of the subgroups formed by such a division will be equal to the posterior probabilities (i.e., $p(S|a_j)$, $p(F|a_j)$ and $p(S|a_j)$, $p(F|a_j)$) generated by Bayes theorem. Subsequent subdivisions in MCR analysis will be based on attributes whose MCR's are computed on the basis of different subgroups. As such, the base rates of these subgroups may or may not be equal to the a posteriori probabilities of Bayes theorem. This equality will depend on whether or not the quantities $p(a_j|S)$ and $p(a_j|F)$ estimated on the
basis of subgroups are equivalent to the corresponding sample estimates. Interestingly enough the use of Bayes theorem assumes that these estimates are equal.

To explain, suppose that $a_1$ and $a_2$ are attributes used for two successive sample subdivisions in an MCR analysis. The base rates of the four subgroups formed on the basis of attribute $a_2$ ($a_2$ being used for the second subdivision) will be equal to the a posteriori probabilities of individuals obtained using Bayes theorem if and only if $p(a_2 | S) = p(a_2 | S, a_1)$ where $p(a_2 | S, a_1)$ is the conditional joint probability of $a_2$ given $S$ and $a_1$. Similarly $p(a_2 | F)$ must equal $p(a_2 | F, a_1)$. These two functions define the assumption of local (conditional) independence necessary for the use of Bayes theorem (Edwards, 1971).

Strictly speaking, therefore, Bayes theorem can be used to predict parole outcome only if the interactions among variables are of relatively little predictive importance for only then will the assumption of conditional independence be met. MCR analysis on the other hand implicitly assumes that such interactions are important enough to be taken into consideration.

The evidence discussed earlier comparing linear and non-linear prediction methods indicated that these procedures were about equivalent in efficiency. This would suggest that the interactions among variables are of relatively little predictive importance. Beverly (cited by Simon, 1971; p. 151) who specifically investigated this issue computed the 741 first-order interactions among 39 predictors of parole outcome. The number of 'significant'
interactions effects was very close to what would have been expected by chance.

There are also a number of theoretical reasons why linear models (see Dawes & Corrigan, 1974) may be more appropriate for predicting parole outcomes, namely (1) parole prediction variables tend to have a conditionally monotonic relationship to the criteria (i.e., variables can be scaled in such a way that higher values in each predict higher values in the criterion, independent of the values of the remaining variables, (2) the relative weights derived are not affected by 'error' in the criterion—such error does not affect the correlation of predicted scores and true scores but does affect the correlation between predicted values and observed values. (3) error in the measurement of predictors tends to make non-linear relationships between dependent and independent variables more linear.

Although it cannot be concluded that interactions among predictors are of no predictive importance, the above discussion suggests that they may be less relevant than was initially supposed by many investigators. It is suggested that the Bayesian technique may be more efficient than the MCR approach because it is less subject to the overfitting of the latter method which results from attempts to capitalize on interaction effects. The Bayesian approach has the additional advantage—like regression—of generating a continuous probability function which allows for greater flexibility in the designation of risk groups.

In addition to comparing the different regression models with the sequential Bayesian technique, a number of prediction parameters
of each method were systematically varied. The research evidence suggests that as the number of independent variables increases, past a certain optimal point, predictive efficiency decreases for the validation samples although it may increase spuriously for the construction sample (e.g., Reiss, 1951). This would appear to be primarily due to overfitting in the development of the prediction model. In the present study methods were compared with respect to the extent to which this occurred for two sets of predictors.

The order in which the independent variables were entered into the prediction models were based on their association with the criterion. The most significant predictor were entered first followed by successively less important predictors. The predictors were ranked in order of importance on the basis of correlation coefficients in the case of multiple regression analysis and on the basis of MCR (i.e., $|p(a_1|S) - p(a_1|F)|$) in the case of the Bayesian analysis.

In addition to varying the number of predictors, the critical $F$ ratios required for entering and dropping variables from the regression models were systematically altered. Specifically, the 'F to enter' of the forward and stepwise models as well as the 'F to remain' of latter techniques were set at two different significance levels which preliminary analyses indicated were appropriate for comparison purposes.

In the Bayesian analyses, the critical value of the a posteriori probabilities which determines procedure termination was varied. It was felt that the most interesting levels of this critical value
for comparisons with the regression technique would be 1.0 and .90. At the 1.0 level the Bayesian procedures will utilize all predictors and are directly comparable to the general regression model. At this criterion level the Bayesian procedures are in effect non-sequential since the characteristics of the data do not allow the a posteriori probabilities to attain a value of 1.0. As a consequence, the a posteriori probabilities are generated on the basis of different attribute vectors of equal, that is maximum, size. At a .90 critical level, on the other hand, it is very likely that all predictors will not be utilized in all cases.

In addition to evaluating the regression and Bayesian approaches as applied to the sample as a whole, comparisons of these techniques were undertaken with respect to moderator variable subgroups. The rationale of the latter strategy will now be developed.

(iii) Moderator Variable Analysis

Strictly speaking, population heterogeneity can be said to reduce the overall efficiency of a particular prediction method if (a) the term 'heterogeneity' as used here refers to \( n \) different patterns of relationships between predictors and criteria and (b) the relationship patterns defining the \( n \) homogeneous subgroups are such that a greater proportion of the criterion variance is accounted for relative to that of the heterogeneous total group. An increase in predictability for certain subgroups, however, might, to a greater or less extent, be due solely to an increase in the variability of the criterion. If in fact this is the only factor acting to increase predictability,
there will be a loss in predictive efficiency in those subgroups which must necessarily show a corresponding decrease in criterion variance.

An interesting strategy for predicting parole outcome has been proposed by Babst, Gottfredson and Ballard (1968) which would capitalize on this state of affairs. They suggest that one might employ association analysis to derive relatively homogeneous subgroups and then apply a regression procedure to the middle subgroups where the least differentiation occurs. Such a procedure would capitalize on the increased criterion variance of the middle subgroups. At the same time the method would avoid losses in prediction for subgroups with more extreme base rates (i.e., subgroups best predicted by association analysis). Serious questions can be raised here, however, with respect to the reliability with which individuals can be assigned to the various risk categories using association analysis.

Another study conducted by Sampson (1974) is also relevant to the above discussion. Initially this investigator employed a stepwise regression technique and derived a ten group base expectancy table. Following this, a Q-factor analytic method was employed and two principal components were extracted from the intra-subject correlation matrix of predictor variables. All subjects were then plotted in reference to these factors and one cluster was derived on the basis of visual inspection. A subsequent regression analysis of the cluster yielded an increase in predictability compared to the sample as a whole. The author attributed this to the homogeneity of the cluster.
From the data reported by Sampson (1974) it is impossible to determine to what extent such improved predictability in the cluster was a function of an increase in criterion variance. It is clear that only an important overall increase in predictive efficiency would justify utilization of the Q-factor technique in addition to the stepwise procedure.

An alternative to the strategies described above involves the derivation of subgroups on the basis of a moderator variable. Such a technique would improve predictive accuracy if the interaction of predictors and moderator added to the criterion variance explained by the predictors alone. Reiss (1951) has employed a moderator of this type. Separate parole prediction analyses were done using white and negro subgroups for three sets (i.e., n = 4, 7, and 13) of independent variables. At this point a re-analysis of certain portions (i.e., validation samples only) of these data will be undertaken for purposes of more clearly developing the rationale of the present strategy.

Reiss (1951) reported that the failure base rate for the white and negro subsamples were 32.7% and 41.7% respectively from which the sample base rate is computed to be 35.0. Considering only the variance of the criteria, one would expect the most accurate predictions to be made in the Negro subgroup with less efficient prediction in the combined group and still poorer results in the white subgroup. This is exactly what was obtained in all analyses with the % reduction in the error of prediction of the respective groups being: (a) 27.6%, 20.5%, and 17.4% (4 predictors), (b) 27.6%, 13.1%, 12.1% (7 predictors),
and (c) 36.2%, 14.3%, and 13.1% (13 predictors). In all three analyses, it can be seen that the relatively small differences between the % reduction in error for the combined and white groups corresponds to the similarly small differences in their respective base rates (i.e., 32.7% vs. 35.0%). Correspondingly large differences can be observed for the negro subgroup with its higher base rate (i.e., 41.7%).

Although the differences in criterion variance across subgroups would appear to be important with respect to predictive efficiency this does not negate the possibility that an overall gain in accuracy was obtained by dividing the sample. In order to determine whether or not this was the case, a proportionately (i.e., of the sample) weighted average of the % reduction in prediction error was computed for the white and negro groups and compared to that of the overall groups. For the 4, 7, and 13 variable analyses these percentages for the combined and original groups were 20.1 vs. 20.5; 16.1 vs. 13.1 and 20.0 vs. 16.8, respectively. The best prediction was clearly obtained in the four variable analyses where there was no increase—in fact there was a slight decrease—in overall accuracy obtained by splitting the sample. It is only in the 7 and 13 variable analyses, where the error variance of prediction increases, that utilization of the moderator proves worthwhile. In neither of these instances, however, does the accuracy of prediction exceed that attained in the original group using 4 variables. It must be concluded, therefore, that splitting the original sample on the basis of race did not improve overall predictive accuracy.
The Reiss (1951) investigation does not rule out the possibility that moderator variables which are important for parole prediction may yet be found. It is suggested, however, that some rationale (statistical or otherwise) should be employed in the selection of such variables in order to avoid the pitfalls inherent in unsubstantiated approaches.

On the basis of his extensive review of the literature related to the well-known 'clinical vs. statistical' controversy, Sawyer (1966) concluded that the most efficient prediction strategies were those involving a mechanical combination of both clinical and statistical data. In the light of this conclusion, a predictive strategy involving the utilization of a quantified 'clinical' variable as a moderator was chosen for purposes of the present study. The variable considered here was the degree of certainty associated with the prognosis of parole outcome made by Classification Officers. The sample was subdivided into two groups such that the members of one consisted of those individuals whom Classification Officers were willing to designate as either 'good' or 'poor' candidates for parole. The other subgroup consisted of cases who received a 'fair' prognosis from Classification Officers.

It is proposed that the overall criterion variance accounted for in such subgroups will be greater than for the sample as a whole. This hypothesis is based on two assumptions, these being (a) cases which Classification Officers feel more certain about with respect to predicting parole outcome will be predicted more accurately on a statistical basis as compared to the sample as a
whole, and (b) the subgroup about whom Classification Officers feel less certain about will be characterized by greater criterion variance than the entire group. It is suggested that this will also tend to increase predictive efficiency relative to the sample as a whole.

(iv) Summary of Analyses and Evaluation of Their Relative Efficiency

General, forward and stepwise regression models were initially developed using two sets of predictors. As discussed, the latter two models were constructed on the basis of two different significance levels of variable selection. This represents ten regression analyses in all and for each of these two measures of predictive efficiency were computed; these being (a) the R statistic and (b) the MCR of two subgroups formed on the basis of whether or not individual predicted scores are greater than or equal to the mean of the means of the predicted scores of the success and failure subgroups. R of course, is the multiple correlation of the predictors and the criterion while the MCR reflects the degree to which the base rates of the subgroups differ from those of the sample as a whole. It should be pointed out here that a higher MCR would be obtained in the present analysis if more than two subgroups were formed. The purpose of the present investigation, however, was to compare the predictive efficiency of various strategies rather than to develop a model of maximum evaluated efficiency. As such, the less cumbersome procedures involving two as opposed to n subgroups was deemed more appropriate.
The next analysis conducted involved the selection of the most efficient of the 10 regression procedures discussed above and substitution of unit for beta weights in the specification equation. R and the MCR—based on a mean split of predicted scores—were also computed here.

The Bayesian technique was evaluated on the basis of four separate analyses. Each of the two sets of predictors employed in the regression procedures were utilized in two Bayesian analyses which differed from each other with respect to the critical a posteriori probability. Here again R and the MCR were employed as measures of predictive efficiency.

As mentioned, the sample was split into two subgroups for those analyses in which the Classification Officer's prognosis was employed as a moderator variable. The most efficient regression and Bayesian models—as ascertained from the analyses of the whole sample—were then applied to each of these subgroups. In order to evaluate the relative predictive efficiency between subgroups, the multiple correlation coefficient was computed separately on the basis of each. This was done for both the regression and Bayesian procedures. Finally, in order to ascertain whether or not the use of a moderator variable improved overall prediction, the multiple correlation coefficient was computed for each technique on the basis of the combined subgroups.
2. Methodological Points of Importance

(i) The Tukey 'Jackknife' Procedure

Tukey (1969) has proposed a general procedure for reducing sample bias and generating statistical confidence limits. This technique—termed the 'jackknife'—involves five steps when applied to a sample for purposes of computing some numbers; namely,

(a) Division of the sample into r 'pieces' or subgroups stratified with respect to important variables.

(b) Calculating the number $y_{all}$ based on all the data.

(c) Calculating the numbers $y_{(i)}$ (where $i = 1, 2 \ldots r$) based on all $r-1$ pieces of data except the $i$th.

(d) Calculating the pseudovalues $y^{*i}$ where $y^{*i} = r \cdot y_{all} -(r-1)$. $y(i)$

(e) Treating the pseudovalues as if they were a sample and applying student's $t$.

Tukey's (1969) general procedure was adopted for the present study in the following manner. In all analyses the sample or subgroup was divided into 4 pieces stratified with respect to the proportion of successes and failures in each. Prediction of cases in each piece was based on the parameters estimated in the remaining 3 pieces. The evaluation of each prediction strategy therefore involved four separate analyses.

In addition to computing $R$ and the MCR for each procedure as described earlier, 95% confidence intervals of the following form were also derived:
\[
\sqrt{\frac{n}{i=1} \frac{R_i^2 - t \cdot 0.025}{\sqrt{n}}} \leq R_p \leq \sqrt{\frac{n}{i=1} \frac{R_i^2 + t \cdot 0.025}{\sqrt{n}}}
\]

OR

\[
\frac{R_{L}}{s} \leq R_p \leq \frac{R_{U}}{s}
\]

where: \( R_i^2 \) = the squared multiple correlation of piece \( i \).
\( t \cdot 0.025 \) = the value of the \( t \)-statistic \( p \leq 0.025 \), \( df = 3 \)
\( n = 4 \) (pieces)
\( s \) = the standard deviation of the squared multiple correlation coefficient of the 4 pieces
\( R_p \) = the (unknown) population multiple correlation coefficient
\( R_{L} \) = estimated lower bound of the population multiple correlation coefficient
\( R_{U} \) = estimated upper bound of the population multiple correlation coefficient.

As a final point it should be noted here that the \( R \) and MCR statistics computed using the jackknife procedure represent lower bound estimates since they are based on validation subgroups.

(ii) Collection and Preparation of Data

The data base consists of information collected from the personal files of individuals who terminates their status as parolees during the period 1970-1975 inclusive. In total, personal data for 205 parolees was obtained from files made available through the Windsor (140 cases) and London (65 cases) branches of the National Parole Service. The data for each individual consists of 86 items of information (62 dichotomous; 24 continuous), most of which were adopted
from previous studies investigating the prediction of parole outcome (see Appendix C).

Variables were eliminated from this preliminary data set on the basis of an initial screening of extreme endorsement frequencies, missing items, and item-criterion association. Only items which met very liberal standards of significance for both the regression ($r, p \leq .10$) and Bayesian (MCR $\leq .10$) procedures were included in the final data set.

(iii) Operational Definition of Criterion and Moderator Variable

For purposes of the present investigation an individual case was defined as a 'failure' on parole if any one of the following criteria were met: (a) if the offender's parole was revoked or forfeited and he was returned to prison, (b) if there was an outstanding warrant for an indictable offence and/or parole suspension which had not been executed at the time of the present study, (c) if the offender was arrested on an indictable charge which was not dealt with by the time of the present study, (d) if the offender dies during the commission of a crime or from a narcotic drug or alcohol overdose, and (e) if the offender was declared criminally insane. All other parole outcomes were defined as successful.

In utilizing the Classification Officer's prognosis as a moderator variable it was essential that the latter be reliably recorded on the basis of predefined criteria. The following definitions, adopted from Duncan, Chlin, Reiss and Stanton (1953) were utilized for this purpose:
Definition

Prognosis

**good**
- offender has the capacity and disposition to adjust to civilian life and will likely succeed on parole with a minimum of supervision.
- offender has certain limitations in capacity and/or disposition to adjust to civilian life but will likely succeed on parole if circumstances are equitable and supervision is adequate.

**fair**
- offender has certain limitations in capacity and/or disposition to adjust to civilian life but may succeed on parole if circumstances are favourable and supervision is close.

**poor**
- offender lacks both the capacity and disposition to adjust to civilian life and would be an unfit risk for parole under any circumstances.
- offender does not have the disposition to adjust to civilian life, regardless of capacity and will be prone to violate parole in spite of circumstances and supervision.
CHAPTER IV
RESULTS

1. Preliminary Analyses of the Data Base

As described earlier the data base was composed of 205 male parolees of whom 147 (71.7%) were classified as 'successes' and 58 (28.3%) as 'failures' on parole on the basis of pre-specified criterion. The average offender was 34 years of age, had a grade 10 level of education and had been sentenced to 56 months in prison for previous offenses.

Of the 86 initial data (62 dichotomous, 24 continuous) recorded for each individual, 15 (dichotomous) items were unavailable for 5% or more of the subjects and were discarded. Of the remaining 71 items, 17 had endorsement proportions of .90 or greater and were likewise deleted from the pool of potential predictors.

The criterion correlations and MCR's of the 54 remaining variables (41 dichotomous, 13 continuous) were then computed. A final nineteen-item predictor pool was formed by deleting an additional 35 variables for which (1) the criterion correlation was not significant at the .10 level and/or (2) the MCR was not greater than or equal to .10. For a sample size of n = 205, this final step was equivalent to deleting a variable if it accounted for less than 1% of the criterion variance.

The final 19-item pool consisted of 14 dichotomous and 5 continuous variables. In the case of four of the five continuous
variables, the criterion correlations which were computed following
dichotomization at the median exceeded the corresponding continuous
variable criterion correlations. All predictors, including the
fifth continuous variable, were therefore dichotomized for both the
regression and Bayesian analyses. In addition to simplifying the
procedure this selection of variables allowed for a more direct
comparison of the relative importance of predictors across
techniques.

Summary descriptions of the final nineteen item predictor
pool are presented in Table 1. In Table 2 the criterion correlations,
MCR's and rank orders (by technique) of the final 19-item predictor
pool are presented. Inspection of these data reveals that there is
a close correspondence between the regression and Bayesian techniques
with respect to the relative importance of predictors. Some shifting
of the regression rank order might be expected, however, if semi-
partial rather than zero order correlations with the criterion had
been employed.

2. Evaluating the Relative Efficiency of Prediction

Preliminary analyses indicated that it would be appropriate
to set the variable selection parameters of the stepwise and
forward regression models at the p ≤ .05 and p ≤ .50 levels. In
addition all analyses were performed using either 6 or 19 predictor
variables.

In Table 3 the actual regression equations corresponding to
each model generated on the basis of the entire sample are presented.
**TABLE 1**

Summary Descriptions of Final Nineteen Item Predictor Pool

<table>
<thead>
<tr>
<th>Item # (see Appendix C)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
<td>- stability of past residence</td>
</tr>
<tr>
<td>18</td>
<td>- property versus person offence</td>
</tr>
<tr>
<td>69</td>
<td>- favourable versus unfavourable community investigation</td>
</tr>
<tr>
<td>16</td>
<td>- length of time between offences</td>
</tr>
<tr>
<td>12</td>
<td>- violent versus nonviolent offence</td>
</tr>
<tr>
<td>15</td>
<td>- number of nonviolent offences</td>
</tr>
<tr>
<td>75</td>
<td>- first born versus later born</td>
</tr>
<tr>
<td>17</td>
<td>- offence committed alone versus in a group</td>
</tr>
<tr>
<td>60</td>
<td>- time lost in remission versus no time lost</td>
</tr>
<tr>
<td>25</td>
<td>- juvenile offences reported or not reported</td>
</tr>
<tr>
<td>40</td>
<td>- resides with relatives or nonrelatives</td>
</tr>
<tr>
<td>36</td>
<td>- prognosis of classification officer</td>
</tr>
<tr>
<td>27</td>
<td>- time sentenced for previous offences</td>
</tr>
<tr>
<td>33</td>
<td>- parents separated or divorced</td>
</tr>
<tr>
<td>34</td>
<td>- sibling has criminal record</td>
</tr>
<tr>
<td>19</td>
<td>- first versus subsequent parole</td>
</tr>
<tr>
<td>64</td>
<td>- plans to reside with relatives following imprisonment</td>
</tr>
<tr>
<td>58</td>
<td>- training and/or education received during imprisonment</td>
</tr>
<tr>
<td>44</td>
<td>- age at time of parole</td>
</tr>
<tr>
<td>Item # (see Appendix C)</td>
<td>Criterion Correlation</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>48</td>
<td>.25</td>
</tr>
<tr>
<td>18</td>
<td>.21</td>
</tr>
<tr>
<td>69</td>
<td>.19</td>
</tr>
<tr>
<td>16</td>
<td>.16</td>
</tr>
<tr>
<td>12</td>
<td>.15</td>
</tr>
<tr>
<td>15</td>
<td>.15</td>
</tr>
<tr>
<td>75</td>
<td>.13</td>
</tr>
<tr>
<td>17</td>
<td>.12</td>
</tr>
<tr>
<td>60</td>
<td>.11</td>
</tr>
<tr>
<td>26</td>
<td>.11</td>
</tr>
<tr>
<td>40</td>
<td>.11</td>
</tr>
<tr>
<td>86</td>
<td>.11</td>
</tr>
<tr>
<td>27</td>
<td>.11</td>
</tr>
<tr>
<td>33</td>
<td>.11</td>
</tr>
<tr>
<td>34</td>
<td>.11</td>
</tr>
<tr>
<td>19</td>
<td>.11</td>
</tr>
<tr>
<td>64</td>
<td>.10</td>
</tr>
<tr>
<td>58</td>
<td>.09</td>
</tr>
<tr>
<td>44</td>
<td>.09</td>
</tr>
</tbody>
</table>
### TABLE 3

Multiple Regression Equations Generated on the Basis of the Entire Sample

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General, 6 predictors</td>
<td>$y' = .279 + .161 X_{48} + .138 X_{18} + .133 X_{59} - .045 X_{16} - .056 X_{12} - .008 X_{15}$</td>
</tr>
<tr>
<td>2. Forward-Stepwise, $p \leq .05$, 6 predictors</td>
<td>$y' = .222 + .212 X_{48} + .203 X_{18}$</td>
</tr>
<tr>
<td>3. Forward-Stepwise, $p \leq .50$, 6 predictors</td>
<td>$y' = .032 + .162 X_{48} + .193 X_{18} + .137 X_{59}$</td>
</tr>
<tr>
<td>4. General, 19 predictors</td>
<td>$y' = -.799 + .169 X_{48} + .192 X_{18} + .106 X_{59} - .050 X_{16} - .052 X_{12} - .072 X_{15}$ + .114 $X_{75}$ + .142 $X_{17}$ + .030 $X_{50}$ + .099 $X_{26}$ + .024 $X_{40}$ + .037 $X_{85}$ - .114 $X_{27}$ + .120 $X_{33}$ + .044 $X_{34}$ + .075 $X_{19}$ + .038 $X_{54}$ + .021 $X_{58}$ + .151 $X_{44}$</td>
</tr>
<tr>
<td>5. Forward-Stepwise, $p \leq .05$, 19 predictors</td>
<td>$y' = -.174 + .193 X_{48} + .244 X_{18} + .132 X_{75} + .130 X_{17}$</td>
</tr>
<tr>
<td>6. Forward-Stepwise, $p \leq .50$, 19 predictors</td>
<td>$y' = -.748 + .203 X_{48} + .245 X_{18} + .120 X_{59}$ + .122 $X_{75}$ + .128 $X_{17}$ + .091 $X_{26}$ + .109 $X_{27}$ + .101 $X_{33}$ + .143 $X_{44}$</td>
</tr>
</tbody>
</table>
In all analyses the stepwise and forward techniques reduced to identical procedures and are presented in the table in this manner. The simplest prediction model was generated by the forward-stepwise technique \((p \leq .05)\) which incorporated 2 predictors and the most complex was the 19 predictor general regression model. The MCR and multiple correlation coefficients of these models presented in Table 4, suggest that equation incorporating more predictors may be more efficient on construction.

In order to evaluate the relative degree of shrinkage of these models each analysis was repeated over four sample pieces each of which was randomly stratified with respect to the proportion of parole successes and failures. The average (across pieces) MCR and multiple correlation coefficients of these analyses are presented in Table 5. More shrinkage occurred in the case of the 19 variable as compared to the 6 variable models but no procedure emerged as being clearly superior to the others on validation. The least shrinkage, as measured by both the multiple correlation and MCR indices, was found in the case of the unit weighting model. The validation values of the multiple correlation and MCR were also largest for the points score model but not significantly so.

It was proposed that 95% confidence limits for the multiple correlation coefficient could be computed based on the construction and validation coefficients corresponding to each of the four pieces. Due to the degree of variation of the latter across
TABLE 4
MCR and Multiple Correlation Coefficients of Regression Analyses
Computed on the Basis of the Entire Sample

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Multiple R</th>
<th>MCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General, 5 predictors</td>
<td>.32</td>
<td>.30</td>
</tr>
<tr>
<td>2. Forward-Stepwise, p ≤ .05, 5 predictors (2 retained)</td>
<td>.29</td>
<td>.29</td>
</tr>
<tr>
<td>3. Forward-Stepwise, p ≤ .50, 6 predictors (3 retained)</td>
<td>.32</td>
<td>.29</td>
</tr>
<tr>
<td>4. General, 19 predictors</td>
<td>.43</td>
<td>.43</td>
</tr>
<tr>
<td>5. Forward-Stepwise, p ≤ .05, 19 predictors (4 retained)</td>
<td>.35</td>
<td>.36</td>
</tr>
<tr>
<td>6. Forward-Stepwise, p ≤ .50, 19 predictors (9 retained)</td>
<td>.41</td>
<td>.41</td>
</tr>
</tbody>
</table>
TABLE 5
Average MCR and Multiple Correlation Coefficients of Regression Analyses Computed on the Basis of the Four Sample Pieces

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Multiple R</th>
<th></th>
<th>MCR</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Construction</td>
<td>Validation</td>
<td>Construction</td>
<td>Validation</td>
</tr>
<tr>
<td>1. General, 6 predictors</td>
<td>.33</td>
<td>.24</td>
<td>.32</td>
<td>.20</td>
</tr>
<tr>
<td>2. Forward-Stepwise, ( p \leq .05 ), 6 predictors (1.8 retained)*</td>
<td>.29</td>
<td>.25</td>
<td>.26</td>
<td>.23</td>
</tr>
<tr>
<td>3. Forward-Stepwise, ( p \leq .50 ), 6 predictors (3.5 retained)</td>
<td>.33</td>
<td>.26</td>
<td>.29</td>
<td>.20</td>
</tr>
<tr>
<td>4. General, 19 predictors</td>
<td>.46</td>
<td>.28</td>
<td>.43</td>
<td>.23</td>
</tr>
<tr>
<td>5. Forward-Stepwise ( p \leq .05 ), 19 predictors (2.5 retained)</td>
<td>.33</td>
<td>.24</td>
<td>.34</td>
<td>.23</td>
</tr>
<tr>
<td>6. Forward-Stepwise ( p \leq .50 ), 19 predictors (8.5 retained)</td>
<td>.44</td>
<td>.25</td>
<td>.41</td>
<td>.20</td>
</tr>
<tr>
<td>7. Unit Weighting (based on Forward-Stepwise ( p \leq .50 ))</td>
<td>.31</td>
<td>.29</td>
<td>.29</td>
<td>.25</td>
</tr>
</tbody>
</table>

*Average number of predictors retained in model in each piece.
pieces, however, it was considered more meaningful to compute only estimations of the upper limits. These estimates were generated for each model at the 95% significance level and are presented in Table 6. Inspection of these data reveal a pattern similar to that found in Table 5. Specifically, the 19 variable models have greater estimated upper limits than the 6 variable models on construction but such differences diminish to the point of nonsignificance upon validation.

The multiple correlation coefficients and MCR indices were computed for the Bayesian analyses of the entire sample and are presented in Table 7. As in the case of the corresponding regression analyses (see Table 4), models incorporating more predictors appear to be more efficient on construction. In the case of both the 6 and 19 variable models, however, the Bayesian approach appears less efficient than regression on construction.

In order to compare the shrinkage patterns of the Bayesian and regression procedures, Bayesian models were generated for the four sample pieces. The average MCR and multiple correlation coefficients of these analyses are presented in Table 8. As was observed for the regression procedures, more shrinkage occurred in the case of the 19 variable as compared to the 6 variable models. Inspection of the validation coefficients in Table 8 however reveals that the Bayesian procedures were of approximately equivalent efficiency in predicting parole outcome.
<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Estimated Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Construction</td>
</tr>
<tr>
<td>1. General, 6 predictors</td>
<td>$R \leq .37$</td>
</tr>
<tr>
<td>2. Forward-Stepwise, $p \leq .05$, 6 predictors (1.8 retained)</td>
<td>$R \leq .35$</td>
</tr>
<tr>
<td>3. Forward-Stepwise, $p \leq .50$, 6 predictors (3.5 retained)</td>
<td>$R \leq .37$</td>
</tr>
<tr>
<td>4. General, 19 predictors</td>
<td>$R \leq .51$</td>
</tr>
<tr>
<td>5. Forward-Stepwise, $p \leq .05$, 19 predictors (2.5 retained)</td>
<td>$R \leq .40$</td>
</tr>
<tr>
<td>6. Forward-Stepwise, $p \leq .50$, 19 predictors (8.5 retained)</td>
<td>$R \leq .50$</td>
</tr>
</tbody>
</table>
TABLE 7
MCR and Multiple Correlation Coefficients
of Bayesian Analyses Computed on the
Basis of the Entire Sample

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Multiple R</th>
<th>MCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Criterion = 1.0, 6 predictors</td>
<td>.25</td>
<td>.18</td>
</tr>
<tr>
<td>2. Criterion = .90, 6 predictors</td>
<td>.21</td>
<td>.18</td>
</tr>
<tr>
<td>3. Criterion = 1.0, 19 predictors</td>
<td>.31</td>
<td>.30</td>
</tr>
<tr>
<td>4. Criterion = .90, 19 predictors</td>
<td>.28</td>
<td>.24</td>
</tr>
</tbody>
</table>
TABLE 8

Average MCR and Multiple Correlation Coefficients
of Bayesian Analyses Computed on the
Basis of the Four Sample Pieces

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Multiple R</th>
<th></th>
<th>MCR</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Construction</td>
<td>Validation</td>
<td>Construction</td>
<td>Validation</td>
</tr>
<tr>
<td>1. Criterion = 1.0, 6 predictors</td>
<td>.32</td>
<td>.28</td>
<td>.29</td>
<td>.20</td>
</tr>
<tr>
<td>2. Criterion = .90, 6 predictors</td>
<td>.30</td>
<td>.27</td>
<td>.28</td>
<td>.21</td>
</tr>
<tr>
<td>3. Criterion = 1.0, 19 predictors</td>
<td>.38</td>
<td>.29</td>
<td>.33</td>
<td>.22</td>
</tr>
<tr>
<td>4. Criterion = .90, 19 predictors</td>
<td>.37</td>
<td>.28</td>
<td>.33</td>
<td>.23</td>
</tr>
</tbody>
</table>
Comparison of the data presented in Table 8 for the Bayesian analyses and the corresponding statistics of the regression procedures (see Table 5) reveals that the former technique exhibited less shrinkage than the least method on validation. Since the Bayesian predictions are generally not as accurate as those of the regression procedures on construction, however, the two techniques performed about equally well upon validation.

The estimated upper limits (p >= .95) of the multiple correlation coefficients associated with each of the Bayesian analyses were computed and are presented in Table 9. As was found in the case of the regression analyses (see Table 5), the 19 variable models have greater estimated upper limits than the 6 variable models on construction but these differences are not discernable on validation. In addition, comparison of the regression and Bayesian validation upper limits indicates there is little variation with respect to this statistic across techniques.

For purposes of the moderator variable analyses, the sample was split into two subgroups. One of these groups (i.e., Group 1) consisted of those parolees who received either a 'poor' or 'good' prognosis from Classification Officers while the other was made up of individuals who received a 'fair' prognosis (i.e., Group 2). Group 1 consisted of 121 subjects of whom 93 (76%) were parole successes while 29 (24%) were failures. Group 2 was composed of 55 (65%) successes and 29 (24%) failures for a total of 84 parolees.
<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Estimated Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Construction</td>
</tr>
<tr>
<td>1. Criterion = 1.0, 6 predictors</td>
<td>$R \leq .37$</td>
</tr>
<tr>
<td>2. Criterion = .90, 6 predictors</td>
<td>$R \leq .35$</td>
</tr>
<tr>
<td>3. Criterion = 1.0, 19 predictors</td>
<td>$R \leq .43$</td>
</tr>
<tr>
<td>4. Criterion = .90, 19 predictors</td>
<td>$R \leq .40$</td>
</tr>
</tbody>
</table>
The results of the analyses described to this point indicated that there was little difference in the efficiency of 6 and 19 predictor models in the case of both the regression and Bayesian validation procedures. In the light of this observation—and the fact that there were fewer subjects in each of the two moderator subgroups as compared to the sample as a whole—the most efficient 6 variable regression and Bayesian models were selected for purposes of the moderator variable analyses. Specifically, the forward-stepwise regression ($p \geq .50$) and Bayesian (criterion = 1.0) models were chosen on the basis of the magnitude of the associated multiple correlation coefficients.

The average multiple correlation coefficients computed on the basis of two sets of four subgroup pieces are presented in Table 10. For comparison purposes the multiple correlation coefficients for the combined subgroups as well as for the total sample analyses (taken from Tables 5 and 8) were also included in the table. As expected, more efficient prediction was obtained in Group 1 as compared to the sample as a whole. In the case of both the regression and Bayesian procedures, however, predictions for Group 2 were less accurate than those obtained for the total sample despite an increase in criterion variance. An inspection of Table 10 also reveals that there was a slight but nonsignificant overall increase in predictive efficiency obtained through the use of the moderator variable for both methods.
### TABLE 10

Average Multiple Correlation Coefficients of Moderator Variable

Analyses Computed on the Basis of Two Sets of Four Sub-group Pieces

<table>
<thead>
<tr>
<th>Group</th>
<th>Model</th>
<th>Multiple R Construction</th>
<th>Multiple R Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Regression</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Forward-Stepwise, p</td>
<td>.50</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.32</td>
</tr>
<tr>
<td>2</td>
<td>Forward-Stepwise, p</td>
<td>.50</td>
<td>.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.19</td>
</tr>
<tr>
<td>1 &amp; 2*</td>
<td>Forward-Stepwise, p</td>
<td>.50</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.27</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample</td>
<td>Forward-Stepwise, p</td>
<td>.50</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.26</td>
</tr>
<tr>
<td><strong>Bayesian</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Criterion = 1.0</td>
<td></td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.34</td>
</tr>
<tr>
<td>2</td>
<td>Criterion = 1.0</td>
<td></td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.22</td>
</tr>
<tr>
<td>1 &amp; 2*</td>
<td>Criterion = 1.0</td>
<td></td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.30</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample</td>
<td>Criterion = 1.0</td>
<td></td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.28</td>
</tr>
</tbody>
</table>

*Coefficients for the combined group were computed as a weighted average of groups 1 and 2.
As in previous analyses the Bayesian procedure was characterized by less shrinkage than the regression technique and was consistently more efficient with respect to prediction in validation. Although this difference in efficiency was reliably found in several analyses, it was not large enough to be considered significant.
CHAPTER V
DISCUSSION

1. Conclusions Regarding the Relative Efficiency of Prediction

The most important overall conclusion of the present investigation is that there were no significant differences between any of the prediction strategies studied. That is, the different regression, Bayesian, and moderate variable procedures all performed about equally well in predicting parole outcome. This finding is consistent with those of other investigations described earlier in which the efficiency of different prediction techniques has been evaluated. Simon (1971), for example, found that the multiple R's computed for several regression analyses ranged from .37 to .55 on construction and from .13 to .18 on validation. The multiple R's for the regression analyses of the present study ranged from .29 to .46 on construction and from .24 to .28 on validation. The slightly higher validation coefficients found in the present study may be attributed to the manner in which the final pool of predictors was derived. Specifically, the procedure whereby thirty-five variables were eliminated from the final group of predictors on the basis of sample estimated predictor-criterion associations appears to have capitalized on chance factors to some extent. The fact that the range of the construction multiple correlation coefficients was lower than that found in previous studies would suggest, however, that overfitting resulting from the variable selection procedure was relatively minor.
The Bayesian analyses which were conducted yielded multiple $R^2$s ranging from .30 to .38 and MCR's between .25 and .33 on construction. Upon validation, the range of these statistics was .28 to .29 and .20 to .23 respectively. By comparison, Simon's (1971) predictive attribute analyses, configuration analysis and MCR analysis produced MCR's on construction ranging between .25 and .45 which dropped to between .17 and .31 on validation. The Bayesian procedure would, therefore, appear to be neither more nor less efficient than prediction techniques which have been previously investigated.

The degree of predictive accuracy obtained in the present study was admittedly rather low. This finding can be attributed to several factors. First of all it should be pointed out that the sample used in the present study consisted of individuals who had been granted terms on parole by the National Parole Board. This selection of prison inmates would result in a restriction of the criterion variance relative to that of the inmate population as a whole. A restriction of this type would of course reduce the magnitude of the multiple correlation coefficient associated with a particular prediction procedure. It should also be pointed out here that prediction studies utilizing select inmate samples such as the present one will yield information relevant to decisions made by parole officers. Such information, however, can be expected to be less selective compared to that which would be associated with predictive devices developed on the basis
of the entire innate population. The difference in selectivity here is of course a direct function of the accuracy of decisions made by National Parole Board.

A second factor limiting the accuracy of predicting parole outcome is a statistical one related to the fact that predicted criterion scores are continuous and actual criterion scores are dichotomous. Specifically, because the distributions of these two sets of scores cannot possibly have the same shape, the point-biserial coefficient relating the two sets has a maximum value of about .80 (Morrally, 1967; p. 132) if the base rate is .50. In the present study, if the predicted scores are assumed to be normally distributed, the maximum point-biserial (i.e., multiple) correlation attainable by any prediction method is only about .74 since the sample base rate of success is .72. In the light of this observation the results of the present study appear more promising than would be the case if multiple R's of 1.0 were theoretically attainable.

Finally, the rather low level of predictive efficiency obtained here and in other studies must be attributed, in part, to the numerous sources of error inherent in the parole prediction paradigm. (The presence of such error in the present analyses was made very apparent by the fact that a three variable points score system performed as well or better than the much-more sophisticated regression and Bayesian strategies). It should be pointed out, however, that the Bayesian technique may
be somewhat less subject to the effects of such error than the regression procedure. This is suggested by the following trends in the data: (1) the Bayesian procedures were less efficient than the regression procedures on construction; (2) the Bayesian procedures exhibited less shrinkage than the regression procedures on validation, and (3) the Bayesian procedures were more efficient than the regression strategies on validation. Like the simple points score system, the Bayesian procedure may therefore be less vulnerable to overfitting on construction than the regression procedures.

On the whole, however, the differences in predictive efficiency between the regression and Bayesian techniques were not substantial. The need for further research investigating these differences within the context of the parole prediction paradigm would therefore appear questionable. Other prediction or decision making problems, however, may raise more interesting points of comparison between the two methods. In situations where more accurate predictions are possible, the degree and type of inter-relationships among the predictors may prove to be of importance. More specifically, since the regression technique takes these relationships into account (via semi-partial correlations and suppressor effects) while the Bayesian approach does not, the latter may in fact be at a disadvantage under certain conditions. Another interesting research question relates to the degree and type of relationship between predictors and the criterion. Both
the regression and Bayesian procedures assume that predictor variables are linearly related to the criterion. In circumstances where the interactions among predictor variables are of predictive importance, the extent to which violation of this assumption differentially affects predictive efficiency could be explored.

2. Implications of Results for Parole Decision Policy

In the present study the most efficient strategy (on validation) for predicting parole outcome was the Bayesian moderator variable (criterion = 1.0) procedure. Using this method, however, only 9% of the variance of the criterion was accounted for. Despite the obvious inefficiency of this approach, it still represents an improvement in accuracy relative to the predictions made by Classification Officers which accounted for only 1.2% of the criterion variance.

In light of the above it is suggested that prediction tables should be one of many sources of information considered in the process of deciding whether or not to parole an individual from prison or in determining the type of parole supervision which he receives. The utilization of this additional source of information could only improve the selectivity of the decisions made. As noted above the results of the present investigation would suggest that Bayesian based prediction tables may be slightly more efficient than tables generated using regression which are presently utilized in the United States. It is suggested,
therefore, that a Bayesian rather than regression strategy be employed in the development of such prediction devices for use within the Canadian parole system. As an alternative to the approach based on Bayes theorem which was employed in the present study, maximum likelihood estimates of regression weights (e.g., see Sprent, 1969) could be employed here.

The results of the present study would also suggest that little is to be gained by constructing prediction tables based on numerous variables. The utilization of relatively few well chosen independent variables would appear to be the best tact.

In the present investigation, the following three predictors were of importance in all of the Bayesian models and in four of the six regression models: (1) Reported that offender did not live at the same residence for two years preceding offence in question vs. Reported that offender lives at same residence for this period; (2) Offence in question was a property offence vs. Offence was a person offence (if both person and property offences then scored property offence) and; (3) Community investigation not requested or requested but not favourable vs. Community investigation requested and favourable or indifferent. It is recommended that these predictors be given particular attention in any attempt to develop Canadian parole prediction tables.
<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>The Parole Act</td>
<td>68</td>
</tr>
<tr>
<td>B</td>
<td>The Parole Agreement</td>
<td>74</td>
</tr>
<tr>
<td>C</td>
<td>Data Collection Form</td>
<td>75</td>
</tr>
<tr>
<td>D</td>
<td>An Introduction to Multiple Regression</td>
<td>85</td>
</tr>
<tr>
<td>E</td>
<td>Regression and Bayesian Piece Specific Multiple Correlation Coefficients, MCR's and Predictor Variables</td>
<td>93</td>
</tr>
</tbody>
</table>
APPENDIX A

The Parole Act
Appendix

PAROLE ACT

An Act to provide the Conditional Liberation of Persons Undergoing Sentences of Imprisonment
(Assented to 6th September, 1958)

HER MAJESTY, by and with the advice and consent of the Senate and House of Commons of Canada, enacts as follows:

Short Title

1. This Act may be cited as the Parole Act.

Interpretation

2. In this Act,

(a) "Board" means the National Parole Board established by this Act;

(b) "inmate" means a person who has been convicted of an offence under an Act of the Parliament of Canada and is under sentence of imprisonment for that offence, but does not include a child within the meaning of the Juvenile Delinquents Act who is under sentence of imprisonment for an offence known as a delinquency;

(c) "magistrate" means a justice or a magistrate as defined in the Criminal Code;

(d) "parole" means authority granted under this Act to an inmate to be at large during his term of imprisonment;

(e) "paroled inmate" means a person to whom parole has been granted;

(f) "parole supervisor" means a person appointed by the Board to guide and supervise a paroled inmate; and

(g) "regulations" means regulations made by order of the Governor in Council.

Board Established

3. (1) There shall be a board, to be known as the National Parole Board, consisting of not less than three and not more than five members to be appointed by the Governor in Council to hold office during good behaviour for a period not exceeding ten years.

(2) The Governor in Council shall designate one of the members to be Chairman and one to be Vice-Chairman.
Appendix

(3) The Governor in Council may appoint a temporary substitute member to act as a member in the event that a member is absent or unable to act.

(4) A majority of the members constitutes a quorum, and a vacancy on the Board does not impair the right of the remaining members to act.

(5) The Board may, with the approval of the Governor in Council, make rules for the conduct of its proceedings and the performance of its duties and functions under this Act.

(6) The head office of the Board shall be at Ottawa, but meetings of the Board may be held at such other places as the Board determines.

(7) The Board shall have an official seal.

Remuneration.

4. (1) Each member of the Board shall be paid such remuneration for his services as is fixed by the Governor in Council, and is entitled to be paid reasonable travelling and living expenses incurred by him while absent from his ordinary place of residence in the course of his duties.

(2) The officers, clerks and employees necessary for the proper conduct of the business of the Board shall be appointed in accordance with the provisions of the Civil Service Act.

(3) The Chairman is the chief executive officer of the Board and has supervision over and direction of the work and the staff of the Board.

Powers and Duties of Board.

5. Subject to this Act and the Prisons and Reformatories Act, the Board has exclusive jurisdiction and absolute discretion to grant, refuse to grant or revoke parole.

6. (1) The Board shall at the times prescribed by the regulations

(a) review the case of every inmate serving a sentence of imprisonment of two years or more, whether or not an application has been made by or on behalf of the inmate, and

(b) review such cases of inmates serving a sentence of imprisonment of less than two years as are prescribed by the regulations, upon application by or on behalf of the inmate.

(2) Upon reviewing the case of an inmate as required by subsection (1) the Board shall decide whether or not to grant parole.

Regulations.

7. The Governor in Council may make regulations prescribing

(a) the portion of the terms of imprisonment that inmates shall serve before parole may be granted,

(b) the times when the Board shall review cases of inmates serving sentences of imprisonment, and

(c) the class of cases of inmates serving a sentence of imprisonment of less than two years that shall be reviewed by the Board upon application.

8. The Board may

(a) grant parole to an inmate if the Board considers that the inmate has derived the maximum benefit from imprisonment and that the reform and rehabilitation of the inmate will be aided by the grant of parole;

(b) grant parole subject to any terms or conditions it considers desirable;

(c) provide for the guidance and supervision of paroled inmates for such period as the Board considers desirable; and

(d) revoke parole in its discretion.

9. The Board, in considering whether parole should be granted or revoked, is not required to grant a personal interview to the inmate or to any person on his behalf.

10. Where the Board grants parole it shall issue a parole certificate, under the seal of the Board, in such form as the Board prescribes, and shall deliver it or cause it to be delivered to the inmate and a copy to the parole supervisor, if any.

11. (1) The sentence of a paroled inmate shall, while the parole remains unrevoked and unforfeited, be deemed to continue in force until the expiration thereof according to law.

(2) Until a parole is revoked, forfeited or suspended the inmate is not liable to be imprisoned by reason of his sentence, and he shall be allowed to go and remain at large according to the terms and conditions of the parole and subject to the provisions of this Act.

Suspension of Parole.

12. (1) A member of the Board or any person designated by the Board may, by a warrant in writing signed by him, suspend any parole and authorize the apprehension of a paroled inmate whenever he is satisfied that the arrest of the inmate is necessary or desirable in order to prevent a breach of any term or condition of the parole.

(2) A paroled inmate apprehended under a warrant issued under this section shall be brought as soon as conveniently may be
before a magistrate, and the magistrate shall remand the inmate in custody until the Board cancels the suspension or revokes the parole.

(3) The Board shall forthwith after a remand by a magistrate under subsection (2) review the case and shall either cancel the suspension or revoke the parole.

(4) An inmate who is in custody by virtue of this section shall be deemed to be serving his sentence.

**Forfeiture of Parole**

Forfeiture. 13. If a paroled inmate is convicted of an indictable offence, committed after the grant of parole and punishable by imprisonment for a term of two years or more, his parole is thereby forthwith forfeited.

**Apprehension upon Revocation or Forfeiture of Parole**

Apprehension. 14. (1) If any parole is revoked or forfeited, the Board may, by warrant under the seal of the Board, authorize the apprehension of the paroled inmate.

(2) A paroled inmate apprehended under a warrant issued under this section, shall be brought as soon as conveniently may be before a magistrate, and the magistrate shall thereupon make out his warrant under his hand and seal for the recommitment of the inmate as provided in this Act.

**Execution of Warrant**

Warrants for apprehension. 15. A warrant issued under section 12 or 14 shall be executed by any peace officer to whom it is given in any part of Canada, and has the same force and effect in all parts of Canada as if it had been originally issued or subsequently endorsed by a magistrate or other lawful authority having jurisdiction in the place where it is executed.

**Recommitment of Inmate**

Place of recommitment. 16. (1) Where the parole granted to an inmate has been revoked, he shall be recommitted to the place of confinement to which he was originally committed to serve the sentence in respect of which he was granted parole, to serve the portion of his original term of imprisonment that remained unexpired at the time his parole was granted.

(2) Where a paroled inmate, upon revocation of his parole, is apprehended at a place not within the territorial division to which he was originally committed, he shall be committed to the corresponding place of confinement for the territorial division within which he was apprehended, to serve the portion of his original term of imprisonment that remained unexpired at the time his parole was granted.

17. (1) When any parole is forfeited by conviction of an indictable offence the paroled inmate shall undergo a term of imprisonment equal to the portion of the term to which he was originally sentenced that remained unexpired at the time his parole was granted plus the term, if any, to which he is sentenced upon conviction for the offence.

(2) The term of imprisonment prescribed by subsection (1) shall be served as follows:

(a) in a penitentiary, if the original sentence in respect of which he was granted parole was to a penitentiary;

(b) in a penitentiary, if the total term of imprisonment prescribed by subsection (1) is for a period of two years or more; and

(c) in the place of confinement to which he was originally committed to serve the sentence in respect of which he was granted parole, if that place of confinement was not a penitentiary and the term of imprisonment prescribed by subsection (1) is less than two years.

(3) Where a paroled inmate is, after the expiration of his parole, convicted of an indictable offence committed during the period when his parole was in effect, the parole shall be deemed to have been forfeited on the day on which the offence was committed, and the provisions of this Act respecting imprisonment upon forfeiture of parole apply mutatis mutandis.

**Additional Jurisdiction**

18. (1) The Board may, upon application therefor and subject to regulations, revoke or suspend any sentence of whipping or any order made under the Criminal Code prohibiting any person from operating a motor vehicle.

(2) The Board shall, when so directed by the Minister of Justice, make any investigation or inquiry desired by the Minister in connection with any request made to the Minister for the exercise of the royal prerogative of mercy.

**Miscellaneous**

19. An order, warrant or decision made or issued under this Act is not subject to appeal or review to or by any court or other authority.
Appendix

Parole Regulations

P.C. 1960 - 681
(as amended)

REGULATIONS MADE UNDER THE PAROLE ACT

1. These Regulations may be cited as the Parole Regulations.

2. (1) The portion of the term of imprisonment that an inmate shall ordinarily serve, in the cases mentioned in this subsection, before parole may be granted, is as follows:

   (a) where the sentence of imprisonment is not a sentence of imprisonment for life or a sentence of preventive detention, one-third of the term of imprisonment impressed or four years, whichever is the lesser, but in the case of a sentence of imprisonment of two years or more to a federal penal institution, at least nine months;

   (b) where the sentence of imprisonment is for life but not a sentence of preventive detention or a sentence of life imprisonment to which a sentence of death has been commuted, seven years.

(2) Notwithstanding subsection (1), where in the opinion of the Board special circumstances exist, the Board may grant parole to an inmate before he has served the portion of his sentence of imprisonment required under subsection (1) to have been served before a parole may be granted.

(3) A person who is serving a sentence of imprisonment to which a sentence of death has been commuted, shall serve the entire term of the sentence of imprisonment unless, upon the recommendation of the Board, the Governor in Council otherwise directs.

(4) The Board shall not recommend a parole, in a case coming within subsection (3), until at least ten years of the term of the imprisonment have been served.

3. (1) In the case of every inmate serving a sentence of imprisonment of two years or more, the Board shall

   (a) consider the case of the inmate as soon as possible after the inmate has been admitted to a prison, and in any event within six months thereof, and fix a date for his parole review;

   (b) review the case of the inmate in order to decide whether or not to grant or recommend parole and, if parole is to be granted, the date upon which the parole is to commence, on or before

   (i) the date fixed for the parole review pursuant to paragraph (a), or
Appendix

(ii) the last day of the relevant portion of the term of imprisonment referred to in subsection (1) of section 2, whichever is the earlier; and

(c) where the Board, upon reviewing the case of an inmate pursuant to paragraph (b) does not at that time grant or recommend parole to the inmate, continue to review the case of the inmate at least once during every two years following the date the case was previously reviewed until parole is granted or the sentence of the inmate is satisfied.

(2) Where an application for parole is made by or on behalf of an inmate who is serving a sentence of imprisonment of less than two years, the case shall be reviewed upon completion of all inquiries that the Board considers necessary but, in any event, not later than four months after the application is received by the Board.

(3) Nothing in this section shall be construed as limiting the authority of the Board to review the case of an inmate at any time during his term of imprisonment.

4. (1) Where the Board receives an application to suspend or revoke a sentence of whipping, the Board shall

(a) determine forthwith if the sentence should be suspended pending further investigation and, if it was so determined, issue an order accordingly;

(b) conduct such investigation as appears to be warranted in the circumstances; and

(c) as soon as possible after completing the investigation, if any, referred to in paragraph (b)

(i) revoke the sentence,

(ii) refuse to revoke the sentence,

(iii) suspend the sentence for any period the Board may deem applicable,

(iv) refuse to suspend the sentence, or

(v) cancel the order of suspension, if any, made pursuant to paragraph (a).

(2) An order of suspension made pursuant to subsection (1) expires ten days before the expiration of any term of imprisonment to which the convicted person, to whom the sentence of whipping relates, has been sentenced unless, before that day, the Board revokes the sentence of whipping.

5. Where the Board receives an application to suspend or revoke an order made under the Criminal Code prohibiting a person from operating a motor vehicle, the Board shall

(a) conduct as quickly as possible such investigation as appears to be warranted in the circumstances; and

(b) determine as soon as possible if the order should be suspended or revoked and, if it so decides, issue an order accordingly.

6. Where the Board suspends or revokes an order made under the Criminal Code prohibiting a person from operating a motor vehicle, the suspension or revocation may be made upon such terms and conditions as the Board considers necessary or desirable.
APPENDIX B

The Parole Agreement
PAROLE AGREEMENT

I clearly understand that I am still serving the sentence imposed but I am being granted parole to permit me to resume my activities as a citizen at large in the community, under supervision. Therefore, in consideration of parole being granted to me, I solemnly agree:

1. To remain, until the expiry of my sentence, under the authority of the National Parole Service Regional Representative in

2. To forthwith proceed directly to ............................................................... and, immediately upon arrival and at least once a month thereafter, to report faithfully to

3. To accept the supervision and assistance of my supervisor.

4. To remain in the immediate area of ...................................................... or as designated by the Regional Representative and, if I have good cause to leave this area, to obtain permission beforehand through my supervisor.

5. To endeavour to maintain steady employment and to report at once to the Regional Representative through my supervisor, any change or termination of employment or any other change of circumstances such as accident or illness.

6. To secure advance approval from the Regional Representative, through my supervisor, if at any time I wish to:
   (a) purchase a motor vehicle;
   (b) incur debts by borrowing money or instalment buying;
   (c) assume additional responsibilities, such as marrying;
   (d) own or carry firearms or other weapons.

7. To abide by all instructions which may be given by my supervisor or by the Regional Representative through my supervisor, and especially with regard to employment, companions, hours, intoxicants, operation of motor vehicles, medical or psychiatric attention, family responsibilities, court obligations.

8. To abide by these special conditions:

9. To forthwith communicate with the Regional Representative, through my supervisor, if I am arrested or questioned by peace officers regarding any offence.

10. To obey the law and fulfill all my legal and social responsibilities.

I have read, or have had read to me, and fully understand and accept the conditions, regulations and restrictions governing my release on parole. I will abide by and conform to them strictly. I also understand that if I violate them in any manner, I may be recommitted.

(name) (number)

Witnessed: ............................................................... (title)

Date of leaving: ..........................................................
APPENDIX C

Data Collection Form
I.D. (  )

Date Parole Terminated:

Recidivism Classification: Failure (  ), Qualified Success (  ), Success (  ).

Reason for Recidivism Classification:

C.O.'s Classification: Poor (  ), Fair (  ), Good (  ).

Reliability: Agree (  ), Disagree (  ).

Item Coding (0) - not available.

(1) or (2) - keyed responses.
( ) 1. L (Lie)
( ) 2. F (Frequency)
( ) 3. D (Depression)
( ) 4. Hy (Hysteria)
( ) 5. Pd (Psychopathic Deviate)
( ) 6. MF (Masculinity-Femininity)
( ) 7. Pa (Paranoia)
( ) 8. Pt (Psychoasthenia)
( ) 9. Sc (Schizophrenia)
( ) 10. Ma (Hypomania)
( ) 11. Si (Social Introversion)

**Criminal History**

( ) 12. Violence reported with regards to offense associated with parole in question (1)

vs.

NO violence reported in this regard (2)

( ) 13. Probation reported (1)

vs.

No probation reported (2)

( ) 14. Number of offenses on record involving violence ( )

( ) 15. Number of offenses on record not involving violence ( )

( ) 16. Length of time between offense associated with parole in question and last previous offense excluding time served in prison—score above median if no previous offense. ( ) months.

( ) 17. Offense in question was reported to be committed alone (1)

vs.

Offense reported to be committed with one or more individuals (2)

( ) 18. Offense in question was a property offense (1)

vs.

Offense was a person offense (2). (If both person and property offenses then score (1)).

( ) 19. Parole in question is a second or subsequent parole for this individual (1)

vs.

First parole (2)

( ) 20. Sex offense(s) reported on record (1)

vs.

Sex offense not reported (2)
21. Last police report prior to offense in question was unfavorable (1)
   vs.
   Police report favourable, indifferent or irrelevant (2)
   Reliability: Agree ( ), Disagree ( ).

22. Offender reported to be remanded without bail while awaiting sentence for offense in question (1)
   vs.
   Reported disposition other than remanded without bail (2)

23. Restitution was not ordered with regards to the offense in question (1)
   vs.
   Restitution was ordered (2)

24. Offense in question was reported to be planned (1)
   vs.
   Offense reported to be unplanned (2)

25. Mandatory supervision (1)
   vs.
   Parole (2)

26. Juvenile offenses reported (1)
   vs.
   No report of juvenile offenses (2)

27. Length of time sentenced for previous offenses ( ) months.

28. Aliases reported (1)
   vs.
   No aliases reported (2)

29. Offense(s) involving a weapon reported on record (1)
   vs.
   No weapons reported (2)

   Interpersonal Relationships

30. No reported dependent children (1)
   vs.
   Dependent children reported (including illegitimate) (2)

31. Size of Family ( )

32. Married (including common-law) at time of sentence for offense in question (2)
   vs.
   Not married (1)
( ) 33. Offender's parents reported to be separated or divorced (1) 
   vs.
   Marital status of offender's parents reported to be other 
   than separated or divorced (2) 

( ) 34. Offender reported to have sibling with criminal record (1) 
   vs.
   No report of sibling with criminal record (2) 

( ) 35. Offender's parent(s) reported to have criminal record (1) 
   vs.
   No report that parent(s) has criminal record (2) 

( ) 36. Parental alcohol (and/or other drugs) abuse reported (1) 
   vs.
   No report of parental alcohol and/or drug abuse (2) 

( ) 37. Age at first leaving home or foster home for six months 
   or more. (    ) years. 

( ) 38. One or both parents reported dead (1) 
   vs.
   Both parents reported to be alive (2) 

( ) 39. Psychiatric disorder (including suicide attempts) in 
   immediate family reported (1) 
   vs.
   No report of psychiatric disorder in immediate family (2) 

( ) 40. Offender reported to be not living at home (i.e., with 
   wife or relatives) at time of arrest for offense in 
   question (1) 
   vs.
   Reported as living at home (2) 

( ) 41. No report of communication with outsiders (letters, visits) 
   during the incarceration in question (1) 
   vs.
   Communication reported (2) 

Demographic Data

( ) 42. Formal education—other than within the penal system 
   (    ) years. 

( ) 43. Age at time of first offense on RCMP record (    ) years. 

( ) 44. Age at time of parole (    ) years. 

( ) 45. Large residential city (greater than 15,000) (1) 
   vs.
   Small residential city (less than 15,000) (2)
( ) 46. Offender reported to be experiencing economic problems at time of the offense in question (1) vs. No report of economic problems (2)

Reliability: Agree ( ), Disagree ( )

( ) 47. Not Canadian born (1) vs. Canadian born (2)

( ) 48. Reported that offender did not live at the same residence for two years preceding offense in question (1) vs. Reported that offender lived at same residence for this period (2)

( ) 49. Father's occupation not professional (1) vs. Father's occupation professional (2)

Reliability: Agree ( ), Disagree ( )

Employment Record

( ) 50. Offender reported unemployed at time of the offense in question (1) vs. Offender reported employed (2)

( ) 51. Offender's parent (head of household) reported unemployed at time of the offense in question (1) vs. Parent reported employed or retired (2)

( ) 52. Offender reported not to have worked at least three consecutive months prior to the offense in question (1) vs. Offender reported to have worked at least three consecutive months (2)

( ) 53. Number of jobs held in five years preceding offense in question ( )

( ) 54. Offender reported to be totally self-supporting by legal means for the period six months prior to the offense in question (1) vs. Reported totally supported by self (2)
Institutional Behaviour

( ) 55. Paroled from maximum security institution for offense in question (1)
vs.
Paroled from non-maximum security institution (2)

( ) 56. Length of sentence for offense in question

( ) 57. No report of therapy received while incarcerated for offense in question (1)
vs.
Report of therapy received (2)

( ) 58. No reported training and/or education received while incarcerated for offense in question (1)
vs.
Report of training and/or education (2)

( ) 59. Criminal offense(s) reported to have occurred while incarcerated for offense in question (1)
vs.
No reported criminal offense (2)

( ) 60. Time lost in remission for offense in question (1)
vs.
No time lost (2)

( ) 61. No report of outside representation on behalf of offender while incarcerated for offense in question (1)
vs.
Outside representation reported (2)

( ) 62. Report of escape or escape attempts (1)
vs.
No report of escape or escape attempts (2)

( ) 63. Paroled from Collins Bay or Joyceville (1)
vs.
Paroled from other institution (2)

Release Plans

( ) 64. Report of intention to live alone or with individual(s) other than relatives (1)
vs.
Reported intention to live with relatives (2)

( ) 65. Job or continued education not reported to be definitely available immediately upon release (1)
vs.
Job or continued education reported to be definitely available immediately upon release (2)
( ) 66. Offender is not reported to have plans regarding securing employment or continuing education (1) vs.
   Plans of this nature are reported (2)

( ) 67. Indebtness specifically reported (1) vs.
   No specific report of indebtedness (2)

( ) 68. Report of plans of returning to live in city where offense in question was committed (1) vs.
   Report of plans to return to another city (1)

( ) 69. Community investigation not requested or requested but not favourable (1) vs.
   Community investigation requested and favourable or indifferent (2)

   Reliability: Agree ( ), Disagree ( )

Personal Data

( ) 70. Large physical structure reported (1) vs.
   Large physical structure not reported (2)

   Reliability: Agree ( ), Disagree ( )

( ) 71. No military experience or dishonourable discharge reported (1) vs.
   Military experience reported (2)

( ) 72. Poor health not reported (1) vs.
   Poor health reported (2)

( ) 73. Alcohol abuse reported (1) vs.
   Alcohol abuse not reported (2)

( ) 74. Abuse of drugs (other than alcohol) reported (1) vs.
   No abuse of drugs (other than alcohol) reported (2)

( ) 75. Later born (1) vs.
   First born (2).
( ) 76. Psychiatric disturbance reported by a psychologist or psychiatrist (1)
    vs.
    No report of psychiatric disturbance (2)

( ) 77. Not supervised by John Howard Society (1)
    vs.
    Supervised by John Howard Society (2)

( ) 78. Offender reported to be an illegitimate child (1)
    vs.
    Reported legitimate (2)

( ) 79. Tatoos reported (1)
    vs.
    No tattoos reported (2)

( ) 80. Negro (1)
    vs.
    Other (2)

( ) 81. Number of items not available (    )

Classification Officer's Assessment

( ) 82. Offender reported to verbalize poorly (1)
    vs.
    No report of poor verbalization (2)

    Reliability: Agree ( ), Disagree (    )

( ) 83. Offender's failure to demonstrate 'insight' or 'self-understanding' reported (1)
    vs.
    No report of failure to demonstrate 'insight' or 'self-understanding' (2)

( ) 84. Offender's attitude to authority reported to be unsatisfactory in some respect (1)
    vs.
    No report of unsatisfactory authority attitudes (2)

    Reliability: Agree ( ), Disagree (    )

( ) 85. Classification Officer's prognosis good or fair (1)
    vs.
    Classification Officer's prognosis poor (2)

( ) 86. Classification Officer's prognosis good (1)
    vs.
    Classification Officer's prognosis fair or poor (2)
APPENDIX D

An Introduction to Multiple Regression
Kerlinger (1973) has defined multiple regression as 'a method of analyzing the collective and separate contribution of two or more independent variables, \( X_1 \), to the variation of a dependent variable \( Y \).' (p. 3). The 'method of analysis' referred to here is expressed by the general regression equation which takes the following form,

\[
Y' = a + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k
\]

where: \( Y' \) = the predicted \( Y \) (raw) score
\( a \) = the intercept constant
\( b_1 \ldots b_k \) = the regression coefficients associated with the independent variables \( X_1 \ldots X_k \).

In essence, multiple regression involves the solution of the above equation for \( b_i \) and \( a \) according to the principle of least squares. This solution involves minimizing the quantity,

\[
\sum_{i=1}^{n} (Y_i - Y'_i)^2
\]

where: \( Y_i \) = the set of \( n \) observed dependent variables scores.

\( Y'_i \) = the set of \( n \) predicted dependent variable scores.

which is alternatively referred to as the sum of the squared errors of prediction, the sum of the squares of the deviations or the sum of the square at the residuals. This quantity (\( \text{ss}_{\text{res}} \)) represents that portion of the total sum of squares of the dependent variable which is not explained by the independent variables in the regression.
equation. As such,

\[ SS_t = SS_{\text{reg}} + SS_{\text{res}} \]

where: \( SS_t \) = the total sum of squares of \( Y \)

\( SS_{\text{reg}} \) = the sum of squares of \( Y \) due to regression

\( SS_{\text{res}} \) = the sum of squares of the residuals

Given the above, the squared correlation \( R^2 \) between the predicted scores \( \hat{Y}_i \) and the actual scores \( Y \) can then be defined as

\[ \frac{SS_{\text{reg}}}{SS_{\text{res}}} \]

The statistical significance of the regression equation itself can likewise be determined on the basis of the following F ratio,

\[ F = \frac{SS_{\text{reg}}/df_1}{SS_{\text{res}}/df_2} \cdot \frac{R^2/k}{(1-R^2)/(N-k-1)} \]

where: \( K \) = the number of independent variables.

\( N \) = the number of subjects.

The correlation between \( Y_i \) and \( \hat{Y}_i \) referred to above is actually the multiple correlation \( R_{Y_1.2..k} \) of the \( k \) independent variables and the dependent variable \( Y_i \). (Note the multiple correlation between the \( k \) independent variables and \( Y_i \) is always 1.) If all of the \( k \) independent variables were uncorrelated then,

\[ R^2_{Y_1.2..k} = r^2_{Y_1} + \ldots + r^2_{Y_k} \]

In an applied regression problem, however, the independent variables are almost always correlated and the following equation is more appropriate,
\[ R^2_{y \cdot 12 \ldots k} = r^2_{y \cdot 1} + r^2_{y(2 \cdot 1)} + r^2_{y(k \cdot 12 \ldots k-1)} \]

That is, \( R^2 \) is equal to the squared zero-order correlation of the first independent variable with the dependent variable \( Y \), plus the square of all the subsequent semi-partial correlations at each step partialling from the variable being entered into the equation all the variables that preceded it.

In order to determine whether or not a particular independent variable \( k_1 \) adds significantly to the variance of the dependent variable accounted for by the \( k-1 \) variables already in the equation, one must test for significance of the semi-partial correlation \( r^2_{y(k_1 \cdot 12 \ldots k_2)} \). Since \( r^2_{y(k \cdot 12 \ldots k_2)} \) represents the increment in \( R^2 \) due to the variable \( k_1 \) (i.e., \( R^2_{y \cdot 12 \ldots k_1} - R^2_{y \cdot 12 \ldots k_2} \)) the significance of any variable in predicting the dependent measure can be tested by the \( F \) ratio,

\[
F = \frac{(R^2_{y \cdot 12 \ldots k_1} - R^2_{y \cdot 12 \ldots k_2})}{(k_1-k_2) / (1-R^2_{y \cdot 12 \ldots k_1}) / (N-k_1-1)}
\]

To this point our discussion essentially provides a description of the manner in which one evaluates the contribution of \( k \) independent variables to the variation of a dependent variable, given a solution to the general regression equation. How then is the general equation solved?

It will be recalled that the general equation is of the form,

\[ Y' = a + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k \]

In practice the solution of this equation for \( b_1 \) and \( a \) is attained.
by solving a set of simultaneous equations of the form,

\[ \beta_1 + r_{12} \beta_2 + r_{13} \beta_3 + \cdots + r_{1k} \beta_k = y_1 \]
\[ r_{21} \beta_1 + \beta_2 + r_{23} \beta_3 + \cdots + r_{2k} \beta_k = y_2 \]
\[ r_{31} \beta_1 + r_{32} \beta_2 + \beta_3 + \cdots + r_{3k} \beta_k = y_3 \]
\[ \vdots \]
\[ r_{kl} \beta_1 + r_{k2} \beta_2 + r_{k3} \beta_3 + \cdots + \beta_k = y_k \]

where: \( \beta_j \) = the beta weight of normalized independent variable \( j \).
\( r_{ij} \) = the correlation between normalized independent variables \( i \) and \( j \).
\( r_{yj} \) = the correlation between normalized independent variable \( j \) and dependent variable \( y \).

This set of equations is derived according to the principles of calculus, such that the solution of \( \beta_j \) will minimize the sum of the squares of the deviations or residuals.

In terms of a matrix solution the above set of equations can be represented as follows:

\[
\begin{pmatrix}
    r_{11} & r_{12} & r_{13} & \cdots & r_{1k} \\
    r_{21} & r_{22} & r_{23} & \cdots & r_{2k} \\
    r_{31} & r_{32} & r_{33} & \cdots & r_{3k} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    r_{kl} & r_{k2} & r_{k3} & \cdots & r_{kk}
\end{pmatrix}
\begin{pmatrix}
    \beta_1 \\
    \beta_2 \\
    \beta_3 \\
    \vdots \\
    \beta_k
\end{pmatrix}
= 
\begin{pmatrix}
    r_{y1} \\
    r_{y2} \\
    r_{y3} \\
    \vdots \\
    r_{yk}
\end{pmatrix}

\[ R_{ij} \beta_j = y_j \]
Thus \( R_{ij} \beta_j = \frac{R_{ij}}{y_j} \)

\[ \beta_j = R_{ij}^{-1} R_{ij} \]

where \( R_{ij}^{-1} \) is the inverse of \( R_{ij} \).

The latter equation allows one to solve for the standard partial regression coefficients \( \beta_j \) which can be used to predict the dependent variable on the basis of normalized independent variables; that is,

\[ Z_y = \beta_1 Z_1 + \beta_2 Z_2 + \ldots + \beta_k Z_k \]

The partial regression coefficients \( \beta_j \) of the general regression equation which are used with the data in its original form can be obtained from the \( \beta_j \) weights as follows:

\[ b_j = \beta_j \frac{S_y}{S_j} \]

where: \( b_j = \) the partial regression weight of variable \( X_j \)

\( \beta_j = \) the standard partial regression weight of variable \( X_j \)

\( S_y = \) the standard deviation of the dependent variable \( Y \)

\( S_j = \) the standard deviation of variable \( X_j \)

Once the \( b_j \) weights have been solved for, the intercept constant \( a \) is computed as:

\[ a = \bar{Y} - b_1 \bar{X}_1 - b_2 \bar{X}_2 - \ldots - b_k \bar{X}_k \]

What has been described above is referred to as the 'general model' of regression. The regression equation derived under this model will incorporate all independent variables regardless of the degree to which they account for the variance of the dependent measure. There are a number of alternative regression models, however, which allow one to select that subset of variables which
best predict the criterion. Two such techniques are the forward and stepwise procedures.

In a forward solution, the independent variable with the highest zero-order correlation with the criterion is first entered into the regression equation. Subsequent variables enter the model on the basis of the magnitude of the squared semi-partial correlation with the dependent measure after the partialling of all variables already in the model. Only those variables which attain a pre-specified level of statistical significance—with respect to their ability to predict the criterion—are incorporated into the equation. The contribution of any independent variable \( k_1 \) is tested on the basis of an \( F \) ratio of the form discussed earlier; that is,

\[
F = \frac{(R^2_{y\cdot12\ldots k_1} - R^2_{y\cdot12\ldots k_2})}{(k_1 - k_2)}
\]

\[
(1 = \frac{R^2_{y\cdot12\ldots k_1}}{N - k_1 - 1})
\]

In the forward model variables which enter the regression equation remain in the order in which they were introduced. No allowance is made for the effect of new variables on the usefulness of variables already in the equation. The stepwise model on the other hand is designed to incorporate such effects.

Like the forward model, at each step the stepwise procedure selects that variable which results in the greatest increase in \( R^2 \) provided the \( F \) ratio associated with it exceeds the pre-specified \( F \) for entering a variable. In addition, at each step, \( F \) ratios are also calculated for all variables when they are entered into
the equation last. If this statistic does not exceed a pre-
specified level, the associated variable is dropped from the
equation. The analysis is complete when no independent variable
not in the equation has an 'F to enter' greater than the pre-
specified F for entering, and no variable in the equation has an
'F to remain' smaller than the pre-specified F for remaining.

In theory, the forward and stepwise procedures should generate
more accurate predictive equations than the general technique
since the former should eliminate error associated with the incor-
poration of relatively poor predictors. This notion is supported
by the work of Halinski and Feldt (1970) who compared the forward,
stepwise, and general models (critical F, p < .05). The results
of this investigation indicated that while the stepwise and forward
procedures yielded comparable results, both of these techniques
were superior to the general model.

All of the regression models discussed make certain assumptions
which must be adhered to if one is to make valid inferences from
sample to population. These assumptions are: (i) Y scores are
assumed to be normally distributed at each value of X, (ii) Y
scores are assumed to have equal variances at each value of X,
(iii) the errors of prediction are assumed to be random and normally
distributed and have equal variances at each value of X. As
Kerlinger (1973) points, however, only serious violation of these
assumptions—especially combinations of them—will distort results
(p. 48).
APPENDIX E

Regression and Bayesian Piece Specific Multiple Correlation Coefficients, MCR's and Predictor Variables
<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Piece</th>
<th>Multiple</th>
<th>R</th>
<th>MCR</th>
<th>Predictor Variables (see Appendix C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward-Stepwise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p .05, 6 predictors</td>
<td>1</td>
<td>.29</td>
<td>.17</td>
<td>.31</td>
<td>48, 12, 69, ,</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.21</td>
<td>.37</td>
<td>.19</td>
<td>48, 69, ,</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.36</td>
<td>.04</td>
<td>.25</td>
<td>48, 69, 18, 15,</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.29</td>
<td>.28</td>
<td>.29</td>
<td>48, 69, 18, 16,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward-Stepwise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p .50, 6 predictors</td>
<td>1</td>
<td>.30</td>
<td>.22</td>
<td>.30</td>
<td>48, 12, 69, ,</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.29</td>
<td>.34</td>
<td>.31</td>
<td>48, 69, 18, 15,</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.38</td>
<td>.12</td>
<td>.25</td>
<td>48, 69, 18, 15,</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.32</td>
<td>.30</td>
<td>.32</td>
<td>48, 69, 18, 15, 16,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 predictors</td>
<td>1</td>
<td>.32</td>
<td>.21</td>
<td>.30</td>
<td>48, 12, 69, 18, 15, 16,</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.29</td>
<td>.28</td>
<td>.30</td>
<td>48, 12, 69, 18, 15, 16,</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.38</td>
<td>.14</td>
<td>.39</td>
<td>48, 12, 69, 18, 15, 16,</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.33</td>
<td>.30</td>
<td>.31</td>
<td>48, 12, 69, 18, 15, 16,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward-Stepwise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p .05, 19 predictors</td>
<td>1</td>
<td>.34</td>
<td>.17</td>
<td>.44</td>
<td>48, 12, , 75, 33,</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.26</td>
<td>.35</td>
<td>.24</td>
<td>48, 69, ,</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.40</td>
<td>.04</td>
<td>.41</td>
<td>48, 69, ,</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.29</td>
<td>.28</td>
<td>.29</td>
<td>48, 69, 18, ,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward-Stepwise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p .50, 19 predictors</td>
<td>1</td>
<td>.46</td>
<td>.10</td>
<td>.39</td>
<td>48, 12, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19,</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.37</td>
<td>.38</td>
<td>.37</td>
<td>48, 69, 18, ,</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.50</td>
<td>.12</td>
<td>.55</td>
<td>48, 69, 18, 75, 33, 44, 17, 27, 34, 66,</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.40</td>
<td>.28</td>
<td>.36</td>
<td>48, 69, 18, 75, 17, 26, 34,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>Piece</td>
<td>Multiple R</td>
<td>MCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------------</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C*</th>
<th>V*</th>
<th>C</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.49</td>
<td>.14</td>
<td>.35</td>
<td>.23</td>
</tr>
<tr>
<td>2</td>
<td>.40</td>
<td>.45</td>
<td>.40</td>
<td>.30</td>
</tr>
<tr>
<td>3</td>
<td>.50</td>
<td>.11</td>
<td>.57</td>
<td>.04</td>
</tr>
<tr>
<td>4</td>
<td>.44</td>
<td>.30</td>
<td>.41</td>
<td>.25</td>
</tr>
</tbody>
</table>

Predictor Variables (see Appendix C)

48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58

*C = construction

*V = validation
<table>
<thead>
<tr>
<th>Bayesian Criterion =</th>
<th>Piece</th>
<th>Multiple R Cn</th>
<th>Multiple R Vn</th>
<th>MCR C</th>
<th>MCR V</th>
<th>Predictor Variables (see Appendix C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>0.30</td>
<td>0.31</td>
<td>0.27</td>
<td>0.29</td>
<td>48, 12, 69, 18, 15, 16</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.27</td>
<td>0.35</td>
<td>0.25</td>
<td>0.20</td>
<td>48, 12, 69, 18, 15, 16</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.36</td>
<td>0.14</td>
<td>0.29</td>
<td>0.16</td>
<td>48, 12, 69, 18, 15, 16</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.36</td>
<td>0.29</td>
<td>0.37</td>
<td>0.16</td>
<td>48, 12, 69, 18, 15, 16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>0.26</td>
<td>0.30</td>
<td>0.25</td>
<td>0.32</td>
<td>48, 12, 69, 18, 15, 16</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.27</td>
<td>0.34</td>
<td>0.22</td>
<td>0.20</td>
<td>48, 12, 69, 18, 15, 16</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.36</td>
<td>0.13</td>
<td>0.30</td>
<td>0.18</td>
<td>48, 12, 69, 18, 15, 16</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.31</td>
<td>0.27</td>
<td>0.34</td>
<td>0.16</td>
<td>48, 12, 69, 18, 15, 16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criterion = 1.0, 6 predictors</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.39</td>
<td>0.24</td>
<td>0.34</td>
<td>0.18</td>
<td>48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.34</td>
<td>0.40</td>
<td>0.29</td>
<td>0.29</td>
<td>48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.41</td>
<td>0.24</td>
<td>0.36</td>
<td>0.24</td>
<td>48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.39</td>
<td>0.24</td>
<td>0.34</td>
<td>0.17</td>
<td>48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criterion = 0.9, 19 predictors</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.38</td>
<td>0.26</td>
<td>0.30</td>
<td>0.26</td>
<td>48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.34</td>
<td>0.34</td>
<td>0.32</td>
<td>0.22</td>
<td>48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.40</td>
<td>0.25</td>
<td>0.36</td>
<td>0.24</td>
<td>48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.37</td>
<td>0.25</td>
<td>0.33</td>
<td>0.19</td>
<td>48, 12, 69, 18, 15, 16, 75, 33, 44, 17, 26, 27, 34, 19, 86, 60, 40, 64, 58</td>
<td></td>
</tr>
</tbody>
</table>

\( ^{nc}C = \) construction
\( ^{V}V = \) validation
REFERENCES


Gough, H. G., Wenk, E. A., & Rozynko, V. V. 'Parole outcome as predicted from the CPI, the MMPI, and a Base Expectancy Table.' Journal of Abnormal Psychology, 1965, 70, 432-441.
Grygier, T. 'The effect of social action—current prediction methods and two new models.' *British Journal of Criminology*, 1966, 6, 269-293.

Grygier, T. 'A computer-made device for sentencing decisions—is further counting and thinking really necessary?' *Journal of Research in Crime and Delinquency*, 1969, 6, 199-209.


Hakeem, M. 'Glueck method of parole prediction applied to 1,861 cases of burglars.' *Journal of Criminal Law, Criminology, and Police Science*, 1945, 36, 97-97.


Hart, H. 'Predicting parole success.' *Journal of Criminal Law, Criminology, and Police Science*, 1923, 14, 405-413.


MacNaughton-Smith, P. *Some statistical and other numerical techniques for classifying Individuals*. (London: Her Majesty's Stationery Office, 1965.)


Sampson, A. 'Post-prison success prediction: A preliminary Florida study.' Criminology, 1974, 12, 155-173.


Slovic, P., & Lichtenstein, S. 'Comparison of Bayesian and regression approaches to the study of information processing in judgment.' Organizational Behavior and Human Performance, 1971, 6, 649-744.


Tibbitts, C. 'Success or failure on parole can be predicted.' Journal of Criminal Law, Criminology, and Police Science, 1931, 22, 11-50.

Tibbitts, C. 'Reliability of factors used in predicting success or failure on parole.' Journal of Criminal Law, Criminology, and Police Science, 1932, 22, 844-853.


Vold, B. G. Prediction Methods and Parole: A study of factors involved in the violation or non-violation of parole in a group of Minnesota adult males. Hanover, 1931b.

Vold, B. G. 'Prediction methods applied to problems of classification within institutions.' Journal of Criminal Law, Criminology, and Police Science, 1936, 26, 202-209.
Wold, G. B. 'Comment on "the efficiency of prediction for criminology".' American Journal of Sociology, 1949, 54, 451-452.


Wesman, A. G., & Bennett, G. K. 'Multiple regression versus simple addition of scores in prediction of college grades.' Educational and Psychological Measurement, 1959, 19, 243-246.

Wilkins, L. T. 'Some developments in prediction methodology in applied social research.' British Journal of Sociology, 1955, 6, 348-363.


VITA AUCTORIS

John Murray Syrotuik was born on September 13, 1950 in St. Catharine’s Ontario where he attended elementary and high school. In October, 1973, he graduated from the University of Waterloo with the Bachelor of Arts Degree. In September, 1973, he began his graduate work at the University of Windsor.