# CLASSIFICATION SYSTEMS FOR THE ACCUSED: AN EMPIRICAL ANALYSIS OF WASHINGTON, D.C.

-- FINAL REPORT --

SUBMITTED TO

NATIONAL INSTITUTE OF JUSTICE
U.S. DEPARTMENT OF JUSTICE

113049

SEPTEMBER 1986

TOBORG ASSOCIATES, INC.

113049

# CLASSIFICATION SYSTEMS FOR THE ACCUSED: AN EMPIRICAL ANALYSIS OF WASHINGTON, D.C.

-- FINAL REPORT --

by

Anthony M.J. Yezer, Principal Investigator
Robert P. Trost, Co-Principal Investigator
Mary A. Toborg, Associate Investigator
Carmela Quintos, Research Assistant
Hormoz Hekmatpanah, Research Assistant



September 1986

Toborg Associates, Inc.
1725 K Street, N.W.
Suite 803
Washington, D.C. 20006

This project was supported by Grant Number 84-IJ-CX-0024, awarded to Toborg Associates, Inc., by the National Institute of Justice, U.S. Department of Justice, under the Justice System Improvement Act of 1979. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

## ACKNOWLEDGEMENTS

We would like to thank a number of individuals who assisted us during the course of this study. Particular thanks are due to officials at the National Institute of Justice, including Mr. James K. Stewart, Director; Dr. Richard Laymon, Program Manager for the Classification, Prediction and Methodology Development Program; and Dr. Richard Linster, Director, Center for Crime Control Research.

Additionally, we would like to thank the D.C. Pretrial Services Agency for the use of the data on which the study is based. In this regard both John A. Carver III, Director, and J. Daniel Welsh, former Director of Research, were especially helpful to us.

Much of the theoretical work that underlies the study was done by Lung-Fei Lee, Professor of Economics at the University of Minnesota. Professor Lee was consistently helpful to us throughout the study as we worked to apply his theoretical framework to an operational problem.

Helpful comments were also received from a number of individuals who attended conferences at which preliminary results from the study were presented. These conferences included the annual conferences of NIJ grantees under the Classification, Prediction and Methodology Development Program, held in Sacramento, California, in 1985 and in Denver, Colorado, in 1986; the American Statistical Association's Annual Meeting in 1985 in Las Vegas, Nevada; the Conference on Asymmetric Information Models in Economics at the University of California at Santa Cruz in 1986; and the 1985 Annual Conference on Pretrial Services, held in Lexington, Kentucky.

We would also like to thank Faye Booker-Logan for the administrative and secretarial assistance she provided to the project. Finally, a number of colleagues provided useful comments at various points in the project. All have our thanks for their assistance.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# I. INTRODUCTION

The complex nature of transactions that occur as the criminal justice system deals with accused individuals arises, in part, due to differences in the information and objectives of various participants in the system. Arrested persons, judges, pretrial services' staff, and bondsmen obviously may have divergent values and interests, but they also have access to different information about the likely consequences of the transactions in which they engage. For example, defendants know far more about their actual guilt or innocence and about the probability that they will engage in further criminal activity or fail to appear for trial. The judge and pretrial services officer have less information about the personal characteristics of the accused and are enjoined by law from using some of this information in the decision-making process. However, the judge may have better information about the way in which the justice system is likely to treat defendants. The bondsman appears to have the poorest access to information but is not prevented from using personal characteristics of the accused in deciding on the terms of bond. Clearly, the pretrial treatment of defendants raises complex issues for anyone attempting to estimate statistical models and to create systems for classification of the accused.

There has been great interest in improving the information available to pretrial services' officers and judges in order to improve decision-making in the area of pretrial release. The primary mechanism for achieving improvement is through using detailed "micro" data on subsequent misconduct of released persons to estimate statistical models of the determinants of misconduct (see, for example, the recent econometric studies by Rhodes [1984], Toborg [1984], Sherwood-Fabre [1984], Goldkamp [1981], and Myers [1981].)

The goal of these previous studies has usually been to find objective indicators of pretrial misconduct which can be applied to the population of arrested persons. This is a very difficult statistical task because, based on their various objectives and imperfect information, the actors in the pretrial release phase of the criminal justice system sort out accused persons and provide them with significantly different treatment. An extreme example of this is that some accused are unable to meet release conditions and remain in jail, while others are released on personal recognizance. Clearly, there are sharp differences in the probability of pretrial misconduct due to this differential treatment. These differences will affect the results of any statistical analysis of the data on subsequent conduct of a sample of arrested persons. The pretrial release system cannot be expected to perform the experiment needed to allow inferences based on simple models - that experiment would involve varying release conditions randomly without regard to personal characteristics of the accused.

The problem can be restated in statistical terms as following Trost and Yezer [1985]. Given that the accused receive differential treatment based on their potential for pretrial misconduct, statistical analysis of the determinants of misconduct conducted for a particular subgroup of persons produces results which are conditional on the prior treatment of that group. Estimates based on such conditional models cannot be extended to the entire population of accused persons and may have limited policy relevance compared to the desired unconditional estimates.

In the example above, one could estimate the determinants of pretrial misconduct for persons released on recognizance, but these estimates would be conditional on the selection rule used by judges and pretrial services' officers in making release decisions. The results would not indicate what would happen if the release rule were changed and persons kept in jail were now released. But for policy purposes, there is usually an emphasis on determining precisely what consequences would follow if the release status of various groups were changed - i.e., on results which are obtained from unconditional estimates which hold for all arrested persons.

Evidence of problems in making inferences using micro data on arrested persons is found in arguments for "bail reform" which contend that the fraction of persons given unconditional release who subsequently misbehave is much smaller than the proportion securing release through bondsmen and conclude that unconditional release should be granted to virtually all arrested persons. Obviously, this reasoning confuses marginal and average propensities to misbehave but it also substitutes conditional probabilities, based on selected subgroups, for the unconditional probabilities on which such policy decisions should be made.

This research develops a statistical method for estimating the unconditional probabilities of misconduct for arrested persons using micro data generated by a pretrial release process that includes a variety of different release terms and conditions. The initial statistical approach was suggested in a theoretical paper by Lee [1984], and this research has developed his initial thoughts into a working program which uses maximum likelihood estimation techniques to determine the unconditional parameters of the pretrial misconduct equation. The estimator, which will hereafter be termed the trivariate probit estimator, is then implemented using data on arrested persons obtained from the Washington, D.C. Pretrial Services Agency (PSA).

The results not only demonstrate the feasibility of using the technique, but they also show the necessity of differentiating between conditional and unconditional estimates. This necessity arises because, for a variety of specific models, the unconditional estimates obtained with the trivariate probit are substantially different than those obtained using single-equation probit estimators which produce conditional estimates. Also, the direction of the differences between the results obtained with the various estimators agrees closely with the differences which would be expected based on theory. While these empirical exercises are developed for a particular pretrial system and data set, they do suggest that conditional estimates of pretrial misconduct obtained using single or even bivariate estimation techniques may suffer from serious biases and should be used with considerable caution.

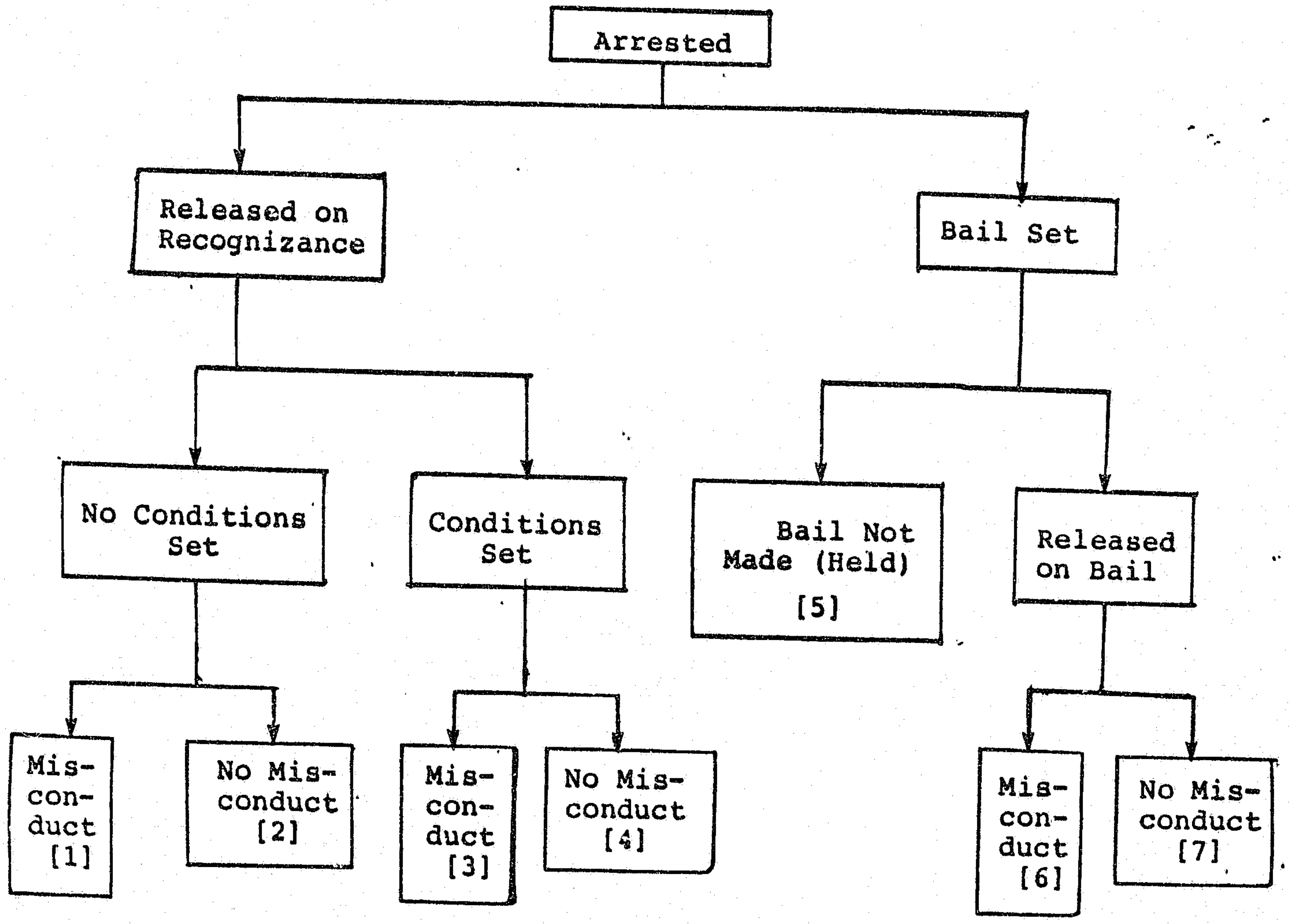
## II. A SIMPLE THEORETICAL VIEW OF PRETRIAL RELEASE AND MISCONDUCT

While the emphasis of this research is on statistical techniques and empirical results, it is important to develop a theoretical approach to the nature of the decision system creating the data being analyzed. Fortunately, the economics of crime literature popularized by Becker and Landes [1974] provides a basis for relating behavior of judges and accused to the general body of microeconomic theory. McFadden [1974] provided the important link between this theory and statistical models of qualitative choice. Taken together, these works, along with subsequent papers offering specific application to criminal justice systems, provide the intellectual foundation for the discussion presented here.

The pretrial release and misconduct process consists of a series of stages in which decisions are made that divide the initial population of accused persons into discrete groups. Figure II-1 presents a simplified diagram of this process. Note that there are seven possible end states [1].......[7] which are separate final groups into which the accused may fall. There are four alternative treatments by the pretrial release system which generate different possibilities and incentives for the accused to make a final decision about pretrial misconduct. At each stage a particular decision maker or makers must make a choice between alternatives which channel the accused toward one path or the other until one of the seven possibilities is realized.

FIGURE II-1

FLOW OF ACCUSED PERSONS THROUGH PRETRIAL RELEASE AND MISCONDUCT



At each stage the decision being made contains two stochastic or random elements. First, individual characteristics of the decision maker make the final choice uncertain. Two judges, if asked to render a decision on pretrial release for identical groups of accused, will not make identical release decisions in all cases simply because judges must differ, however slightly, on the relative importance of the right of the accused to be released vs. the need to ensure appearance or avoid danger to the public. A second stochastic element is the underlying uncertainty regarding the likelihood of pretrial misconduct, should the accused receive a particular type of release.

Following McFadden [1974] in general and Myers' [1981] application to pretrial misconduct, this approach can be applied directly to the release decision of a particular actor, such as a judge who is deciding whether to release on recognizance or set bail. The judge realizes a level of utility, $U_M$, if the accused is freed and engages in pretrial misconduct and a level of utility $U_{NM}$ if the accused is freed and does not commit misconduct. Finally, the judge achieves utility $U_{NF}$ if the accused is not freed - and hence there is no misconduct. The judge must form a conditional expectation of the probability that the accused will engage in misconduct under the following circumstances: release on recognizance, $P_R$, and release on bail, $P_B$.

$P_B$ is the product of the probability of raising bail, $P_{BR}$, and the probability of misconduct conditional on achieving freedom on bail. Now the judge may calculate the expected utility if the accused is released on recognizance,

$U_R=P_RU_M+(1-P_R)U_{NM}$. Expected utility if the bail is set is $U_B=P_BU_M+(1-P_{BR})U_{NF}$. The judge will release the accused on recognizance if $U_R>U_B$. However, the probabilities in the expressions for $U_R$ and $U_B$ are random variables which depend on the personal characteristics of the accused and of the judge forming the expected probability.

Thus, the probability of release on recognizance, $P(U_R>U_B)$ will be a function of the characteristics of the accused and the judge. We write the expected utility if the $i^{th}$ person is released on recognizance as: $U_{Ri}=Z_{Ri}g+e_{Ri}$ and if bail is set as: $U_B=Z_{Bi}g+e_{Bi}$, where $Z_{Xi}$ is a vector of personal charateristics of the $i^{th}$ accused, including criminal justice status and record, g is a vector of parameters, and the e's are continuous variables. In any individual case, the accused is either released on recognizance or bail is set. Let $y_i=1$ indicate that the $i_{th}$ person has bail set. Then the probability of bail can be expressed as $P(y_i=1)$ or as:

$$P(y_i=1)=P(U_B>U_R)=P(Z_{Bi}g+e_{Bi}>Z_{Ri}g+e_{Ri})=P(e_{Bi}-e_{Ri}>g(Z_{Ri}-Z_{Bi})$$
$$=F(g(Z_{Ri}-Z_{Bi}),$$

where F is the distribution function of $e_{Bi}-e_{Ri}$.

In the research reported here, this distribution function will be assumed to be normal and F() will be the cumulative normal or probit. Once a distribution function has been assumed for $(e_{Bi}-e_{Ri})$, the vector of parameters, g's, can be estimated using single-equation techniques, in this case single-equation probit. As noted above, the final disposition of an accused moving through the pretrial release system involves several stages of decision-making. However, the basic economic model underlying each decision is rooted in the expected utility model, and hence this should be recalled when subsequent statistical models are presented below. For example, the decision of an accused to engage in pretrial criminal activity is based on the probability that the expected utility of criminal activity is larger than that if no crime is committed.

## III. PROBLEMS IN PRODUCING INFORMATION ON PRETRIAL MISCONDUCT

The general statistical or econometric problem which makes it difficult to make inferences about the causes and prediction of pretrial misconduct arises due to partial observability of outcomes. This is illustrated in Figure II-1, above, where the tree structure of the process through which the accused flow segments them into different subsamples which are given different treatments. In essence, no controlled experiment is performed with random allocation to pretrial treatment strategies. Therefore, analysis of pretrial misconduct for any subgroup of the accused cannot, in most cases, be used to make inferences about how the general accused population would respond to particular treatment. This is a special case of the general problem of partial observability which has been analyzed recently in the literature. Specifically, the effect of giving treatments to a random sample of accused is not fully observed because part of the sample is excluded from experiencing certain outcomes.

Most recent discussion of the problem of partial observability has been based on the bivariate probit model which has been developed during the last five years in articles by Poirier [1980], Connolly [1983], Farber [1983], Abowd and Farber [1983], Fische, [1981], Danzon and Lillard [1982], Venti and Wise [1982], and Meng and Schmidt [1985]. This sudden and extensive eruption of research which builds upon Zellner and Lee [1965], who worked on the case of full observability, has seen the bivariate probit applied to topics from the outcome of committee voting, through labor negotiations, and decisions to attend college.

The bivariate probit model has two equations, each involving a separate stage of the decision tree and having the following general form:

(III-1) $Y_{1i}* = G_1 + Z_{1i}g_1 + e_{1i}$

$Y_{2i}* = G_2 + Z_{2i}g_2 + e_{2i}$

where $Y_{ji}*$ is the probability of the $j^{th}$ decision, $G_j$ is a constant term, $Z_{ji}$ is a matrix of observed values of independent variables, $g_j$ is a vector of parameters to be estimated, and $e_{ji}$ is an identically and independently distributed random variable. We observe $Y_{ji}=1$ if $Y_{ji}*>0$, otherwise $Y_{ji}=0$ for $j=1,2$. The errors, $e_{ji}$, are assumed to be identically distributed as a standard bivariate normal with correlation $r_{12}$.

In the case of full observability, the values of both the $Y_{ij}$'s are always observed, and the two probit equations can be estimated separately on the entire sample. If $r_{12}$ is not equal to zero, there is an efficiency gain in estimating the equations jointly, but a single equation approach still yields unbiased results. The expected value of $e_{2i}$ equals zero, $E(e_{2i})=0$, because the second decision is observed regardless of the value of $e_{1i}$. The selectivity bias discussed below arises because the second decision is only observed for certain values of $Y_{1i}$ and hence the probability of observing the second decision depends on $e_{1i}$. Then, if $r_{12}$ is nonzero, $E(e_{2i})$ will not be zero either, and an assumption needed for unbiased single equation estimation is violated.

It is important to differentiate cases in which the $Y_{ij}*$'s are generated by joint or simultaneous decisions from those in which the decisions are sequential. This difference is most important for the consequences of partial observability. If the $Y_{ij}$*'s are jointly determined, then they are always generated for each i in the sample and partial observability is literally a data collection problem - although perhaps one that cannot be resolved.

One example of simultaneity is the retirement of a worker from a firm. This involves the joint decisions of the worker and firm but only the outcome, continue working or retire, is observed. If $Y_{1i}$ =1 indicates the worker wishes to continue working and $Y_{2i}$ =1 that the firm wishes the worker to continue, we observe $Y_{2i}=Y_{1i}=1$ as continued work, but the other three possible combinations of the $Y_{ij}$'s are not separately observed. Instead, they are joined in the single observation of retirement. Thus, of four possible outcomes, only one is actually observed and the other three are combined in a single outcome. If there is full information on the decisions made by either the firm or the worker, then the extent of partial information is reduced but not eliminated. If $Y_{2i}$ for the worker is known, then the outcome $Y_{2i}=1$ $Y_{1i}=0$ can be distinguished from the other two cases in which there is a retirement, but $Y_{2i}=0$ $Y_{1i}=0$ and $Y_{2i}=0$ $Y_{1i}=1$ cannot be separated. Alternatively, information on the firm's choice would also leave a different range of partial observability.

If the partial observability arises as a result of sequential decisions such as those in the pretrial release process, there may be a selectivity problem which may be formulated as a bivariate probit estimation problem. In such cases, $Y_{1i}=0$ would result in a failure to observe $Y_{2i}$ so that the separate outcomes $Y_{1i}=0$ $Y_{2i}=1$ and $Y_{1i}=0$ $Y_{2i}=0$ cannot be distinguished. In most cases, the partial observability of sequential behavior is not a data problem. Partial observability arises because the first decision determines whether a second decision is made. For example, a judicial decision to hold an accused person eliminates the possibility of observing the behavior of that individual when released.

Partial observability introduces significant estimation problems. When the first probit equation can be fully observed, estimation by single equation probit is possible but inefficient unless $r_{12}=0$. If the first equation is not fully observed, then the two-equation system must be estimated jointly. In any case, joint estimation is required for the second equation unless $r_{12}=0$ and selectivity bias is eliminated.

The nature of the selectivity bias in the pretrial release system can be illustrated with the simple example developed in the discussion of theory where we reduce the system to two binary decisions. Let $Y_1$ be the judge's release decision with $Y_1=1$ observed if bail is set and $Y_1=0$ for release on recognizance. Allow $Y_2$ to be pretrial misconduct with $Y_2=1$ if there is misconduct and $Y_2=0$ otherwise. The error terms $e_1$ and $e_2$ include the influence of a variety of factors which are difficult to observe and yet may influence the release and misconduct decisions.

It is reasonable to believe that many of the factors in $e_1$ are also in $e_2$. An omitted variable which is positively related to pretrial misconduct will also tend to be positively related to release on bail by judges who wish to deter misconduct. Thus, we expect that the correlation between $e_1$ and $e_2$, $r_{12}$, is likely to be positive. But $e_1$ is also positively associated with the probability of bail being set as seen directly from equation (1), $E(Y^*_{1i}|e_{1i}>0)>0$ which states that the expected value of $Y^*_{1i}$ conditional on $e_{1i}$ being positive is positive.

If we consider estimation of the misconduct equation for the subsample of persons released on recognizance, $Y_{1i}=0$, then the expected value of the error term in the second equation will be negative, $E(e_{2i}|Y_{1i}=0)<0$ because we have oversampled cases in which $e_{1i}<0$, or $E(e_{2i}|e_{1i}<0)<0$. Given that $r_{12}<0$, if $E(e_{1i})<0$ then $E(e_{2i})<0$ and the estimated constant term of the second equation, for misconduct, will be biased downward. This would give the impression that misconduct was less likely among those released on recognizance than one would obtain if the data used for the estimation had been generated by releasing accused persons randomly. Obviously, the danger for policy purposes is that the possibility of misconduct among those forced to post bail if they were released would be underestimated. In addition, the individual coefficient estimates, the other g's, in the second equation may be biased also, but the direction of bias depends on the correlation between the independent variables, Z's, and $e_{1i}$.

## IV. LEE'S METHOD AND MULTI-STAGE SELECTIVITY

Our ability to deal with decisions characterized by partial observability is limited to the bivariate case due to computational difficulties in integrating the multivariate normal distribution. Alternative approaches to estimating multivariate probit probabilities were explored by Lerman and Manski [1981]. They conducted only simple monte carlo experiments and found some success using a method proposed by Clark [1961] and examined by Daganzo [1977]. However, it is not clear theoretically why the Clark approach works, and it has not been used to estimate models with the type of sequential selectivity properties analyzed here.

Of course, the real world does not recognize computational tractability as a limit on complexity, and one could easily argue that three or more levels of decision-making are the rule rather than the exception. Certainly, this is the case with the criminal justice system where arrested persons deal with magistrates, judges, bondsmen, and, of course, finally with their own decisions regarding criminal behavior.

Lee and Maddala [1983] noted the differences in complexity between joint and sequential decisions, which may only be defined for some subpopulation. Lee [1984] has proposed an alternative and computationally more tractable procedure for estimating sequential decision models with censored outcomes. Consider the specification of a discrete choice model with three sequential decision rules (a trivariate model):

(IV-1) $Y_{ji}{}^{*} = Z_{ji}g_{j} - v_{ji}, \quad j = 1, 2, 3$

where we observe $Y_{ji} = 1$ is $Y_{ji}{}^{*} > 0$, else $Y_{ji} = 0$.

If we let X denote a multivariate normal probability, the likelihood function for an observation with total observability is:

(IV-2) $(1-Y_1)(1-Y_2)(1-Y_3)\ X(Y_1=0,Y_2=0,Y_3=0)$
$+Y_1(1-Y_2)(1-Y_3)\ X(Y_1=1,Y_2=0,Y_3=0) + (1-Y_1)Y_2(1-Y_3)\ X(Y_1=0,Y_2=1,Y_3=0)$
$+ (1-Y_1)(1-Y_2)Y_3\ X(Y_1=0,Y_2=0,Y_3=1) + Y_1Y_2(1-Y_3)\ X(Y_1=1,Y_2=1,Y_3=0) +$
$Y_1(1-Y_2)Y_3\ X(Y_1=1,Y_2=0,Y_3=1) + (1-Y_1)Y_2Y_3\ X(Y_1=0,Y_2=1.Y_3=1) + Y_1Y_2Y_3$
$Y_1Y_2Y_3X(Y_1,Y_2=1,Y_3=1)$

If there is partial observability due to sequential decisions, $Y_2$ is only observed if $Y_1=1$ and $Y_3$ is observed only if $Y_2=1$. The likelihood function for an observation is:

(IV-3) $(1-Y_1)\ X(Y_1=0) + Y_1(1-Y_2)\ X(Y_1=1,Y_2=0)$
$+ Y_1Y_2(1-Y_3)\ X(Y_1=1,Y_2=1,Y_3=0) + Y_1Y_2Y_3\ X(Y_1=1,Y_2=1,Y_3=1).$

One method of estimating the parameters $g_j$, $j = 1, 2, 3$, is to maximize the likelihood function given by (IV-3). In this simple three-stage sequential model, the maximization procedure requires the difficult procedure of computing trivariate normal probabilities with numerical methods. If the model contains four stages, then one would have to compute quadruple integrals to maximize the likelihood, five integrals for five stages, etc. One way to avoid this problem is to assume independence among the various decisions, $r_{jk}=0$ all $k>j$, but this yields biased estimates if the independence assumption is false. In the applications to the justice system presented here, the independence assumption is generally inconsistent with the proper operation of the system because individuals are selected for differential treatment based on differences in the likelihood that they will engage in prohibited behaviors in the future.

Lee [1984] has suggested a method which only requires computation of bivariate distributions, no matter how many stages are present in the decision process. However, his method relies on the sequential nature of the decisions and hence will not work for joint decisions. The idea behind Lee's approach is to specify only the marginal distributions $F_j(v_j)$ for all $v_j$ and the bivariate distributions $F_{1k}(v_1,v_k)$ $k>1$. No assumptions are made about higher distributions such as $F_{123}(v_1,v_2,v_3)$, etc., or about the bivariate distributions $F_{jk}(v_j,v_k)$ $k>j>1$. Lee does, however, specify bivariate distributions for the conditional (on $Y_1=1$) random variables $v_j$ and $v_k$ $k>j>1$, after these conditional random variables have been transformed to normality.

For example, consider the three stage sequential model represented by (IV-1). If we specify the underlying distributions $F_1$, $F_{12}$, $F_{13}$, then it follows that:

(IV-4) $F_{j|1}(v_j)=F_{1j}(Z_1g_1,v_j)/F_1(Z_1g_1)$ for $j = 2, 3$.

These conditional variables will have a skewed distribution if correlation between $v_j$ and $v_k$ $(k>j)$ exists, but can be transformed into a normal distribution by:

(IV-5) $v_{j*|1} = X^{-1}(F_{j|1}(v_j))$ $\quad j>1$

where X is the standard normal distribution.

Let $X_2(v_2*, v_3*, r_{23}*)$ be the standard bivariate normal distribution with correlation coefficient $r_{23}*$. A bivariate distribution for the conditional (on $Y_1=1$) random variables $v_2$ and $v_3$ with marginal distributions $F_{2|1}(v_2)$ and $F_{3|1}(v_3)$ can be specified as:

(IV-6) $Q_{23} = X_2(X^{-1}(F_{2|1}(v_2), X^{-1}(F_{3|1}(v_3)); r*_{23})$.

The conditional distributions of $v_j$, $j>2$, in the third and subsequent stages are:

(IV-7) $F_{j|2}(v_2) = Q_{2j}(Z_2g_2,\ e_j)/F_{2|1}(Z_2g_2)$

Note that at the third stage decision, the probability that $Y_3=1$ will be $F_{3|2}(Z_3g_3)$. If one assumes that the underlying distributions of $F_1$, $F_{12}$, and $F_{13}$ are normal (although other distributions are possible), then the likelihood function for an observation will be:

(IV-8) $L=(1-Y_1)(1-X(Z_1g_1)) + Y_1(1-Y_2)(X(Z_1g_1)) - X_2(Z_1g_1,Z_2g_2;r_{12}))$
$+ Y_1Y_2(1-Y_3)(X_2(Z_1g_1,Z_2g_2;r_{12}) - X(Z_1g_1)\ Q_3(Z_2g_2,Z_3g_3,r^*_{23}))$
$+ Y_1Y_2Y_3\ Q_{23}(Z_2g_2,Z_3g_3;\ r^*_{23})\ X(Z_1g_1).$

This approach can be generalized to accommodate four or more sequential stages. Unlike the multivariate normal approach, the likelihood function produced here involves the computation of univariate and bivariate distributions, no matter how many stages are present in the model. Like the multivariate normal approach, Lee's method does allow for the possibility that each decision at an earlier stage can influence the decisions at subsequent stages, i.e., the method allows for the possibility of self-selectivity bias in the observed data.

## V. DATA TO BE USED IN THE EMPIRICAL ANALYSIS

The goals of this research were to implement a multivariate probit estimator using data from the pretrial release system and to test the hypothesis that sequential selectivity effects could have a significant influence on estimation results. This requires comparison of multivariate and single-equation probit estimators. Therefore, the data should be capable of monitoring the progress of arrested persons through a pretrial system at least as complex as that shown in Figure II-1. Fortunately, a high quality micro data set on arrests was available from the Pretrial Services Agency of Washington, D.C. (hereafter PSA), which has a computerized data base on the population of arrested persons.

Given that the objective was the study of pretrial misconduct, it was necessary to select a sampling procedure that would allow the observation of initial arrest followed by the pretrial period during which subsequent arrest or failure to appear might occur. Data on all arrested persons for the mid-1980 to end-of-1982 period were obtained from PSA. The population of all persons arrested in the first half of 1981 was used for this analysis, so that sufficient time was available to observe subsequent pretrial misconduct by the end of 1982. Of course, most of the cases reached final disposition by early 1982. Only cases in which charges were actually filed were considered to be "arrest" instances, and hence instances of "no paper" were dropped as were fugitive warrants from other jurisdictions, and similar highly unusual arrest instances. All arrests were for crimes to be adjudicated in the D.C. Superior Court.

Cases of pretrial arrest were detected by scanning the arrest records during the pretrial period and determining if the individual, identified by a unique police identification number, experienced subsequent arrest, regardless of the type of charge. Failure to appear was based on cases in which a bench warrant was issued.

Obviously, it was possible for the same individual to be arrested several times during the first half of 1981, and hence such individuals could potentially generate several episodes of pretrial misconduct. The sampling technique used here considered only one arrest per person during the six-month period and hence may be said to rely on "person-based" rather than "arrest-based" sampling in which each arrest during the period would generate one pretrial misconduct episode. The problem with using "arrest-based" sampling, as has been the case with prior studies by Myers [1982], Rhodes [1984], and Toborg [1984] is that the probability of being selected for inclusion in the initial stage of the sample at arrest is itself an increasing function of the subsequent selection criteria used to guide pretrial release.

In effect, the sample selectivity problem with "arrest-based" sampling takes the tree in Figure II-1 and makes it a complete circle with instances of pretrial misconduct at the end of the tree generating subsequent observations of arrest at the top of the tree. It is not clear that there is an adequate statistical technique for dealing with this type of circular or simultaneous, multi-stage selection problem.

The consequence of using person-based sampling is that the estimation results are valid for making inferences about the population of arrested persons in a given time period. They would be valid for the population of arrested persons if there were a policy of detaining until trial any person engaging in pretrial misconduct. These sampling and selection issues are quite important but have not been, to our knowledge, given any attention in the previous literature.

The flow of accused through the pretrial release system in Washington, D.C. involves an initial evaluation by PSA which makes a recommendation concerning safety and flight conditions which would be appropriate should the accused be released on recognizance. The judge, usually a hearing commissioner specializing in pretrial release decsisions, then uses the PSA information at arraignment where an initial decision to release on recognizance, set money bail, or hold without bail is made. As a matter of policy, PSA never recommends that money bail be set. PSA's findings of fact concerning the accused may influence the judge. Based on previous observation of the pretrial release system by Toborg [1984], it is likely that the PSA recommendation has a substantial influence on setting release conditions for safety and flight. As a practical matter, judges use setting of money bail as a basic alternative to the use of the conditions recommended by PSA. In the subsequent analysis, we will treat the setting of conditions vs. bail setting as a separate stage in the release process.

Table V-1 contains a glossary of variables commonly used in subsequent empirical analysis. Note that many of the variables,

TABLE V-1

GLOSSARY OF VARIABLES FOR EMPIRICAL ANALYSIS OF PRETRIAL MISCONDUCT

| | |
|---|---|
| Age | Age at arrest in years |
| $Age^2$ | Age squared |
| Bond | Dollar amount, in thousands of dollars, of bond which was set, 0 if Bondset is equal to 0. |
| Bondpost | "Dummy" variable equal to 1 if the variable "Bondset" was equal to 1 and if the data record indicates that bond was posted and equal to 0 otherwise. |
| Bondset | "Dummy" variable equal to 1 if the accused was given an initial release condition that called for percentage bond, cash bond, surety bond, station house bond, and/or a combination of these financial conditions alone or with nonfinancial conditions and equal to 0 otherwise. |
| Confid | "Dummy" variable equal to 1 if most serious charge at arrest is for a confidence crime, fraud, or forgery, and equal to 0 otherwise. |
| Drugs | "Dummy" variable equal to 1 if most serious charge at at arrest is for drug crime and 0 otherwise. |
| Employd | "Dummy" variables equal to 1 if accused was employed at time of arrest and 0 otherwise |
| Excon | Number of prior convictions of accused |
| FTA | "Dummy" variable equal to 1 if the accused failed to appear in a fashion that resulted in the issuance of a bench warrant, and equal to 0 otherwise. |
| Larceny | "Dummy" variable equal to 1 if most serious charge at arrest is larceny and 0 otherwise |
| Male | "Dummy" variable equal to 1 if accused is male, 0 otherwise |
| Miscrim | "Dummy" variable equal to 1 if most serious charge at arrest is for burglary, shoplifting, or similar offense and equal to 0 otherwise. |
| Pendcase | Number of pending charges against accused at time of arrest |

TABLE V-1 CONTINUED

| | |
|---|---|
| Posscrim | "Dummy" variable equal to 1 if most serious charge at arrest was for possession of implements of crime and equal to 0 otherwise. |
| Prosty | "Dummy" variable equal to 1 if most serious charge at arrest was for prostitution and equal to 0 otherwise. |
| Ptarest | "Dummy" variable equal to 1 if the arrested person was arrested subsequently before there was a final disposition of the case or before the end of the data collection period, December 1982, and equal 0 otherwise. |
| Release | "Dummy" variable equal to 1 if the accused was released on either recognizance or with financial conditions and equal to 0 otherwise. |
| Violent | "Dummy" variable equal to 1 if most serious charge at arrest is for a violent crime, murder, rape, or robbery, and 0 otherwise. |
| Weapons | "Dummy" variable equal to 1 if most serious charge at arrest was for a weapons violation and 0 otherwise. |

particularly those reflecting most serious charge at arrest, have been formulated as discrete, zero - one, dummy variables. Of course, the dependent variables in the analysis, such as PTAREST, pretrial arrest, are only observed as discrete zero - one outcomes.

It is also important to recognize that the data used in this analysis are based on a data collection system installed by the D.C. Pretrial Services Agency. The definitions of variables are based on PSA conventions, and the distribution of these variables is based on local conditions in the District of Columbia. Thus, the type of charge is based on the most serious charge at arrest. Clearly, this depends on the type of offenses in the District of Columbia and the charging policies. Finally, the distribution of demographic characteristics of the accused population is also based on the demographic composition of the District of Columbia.

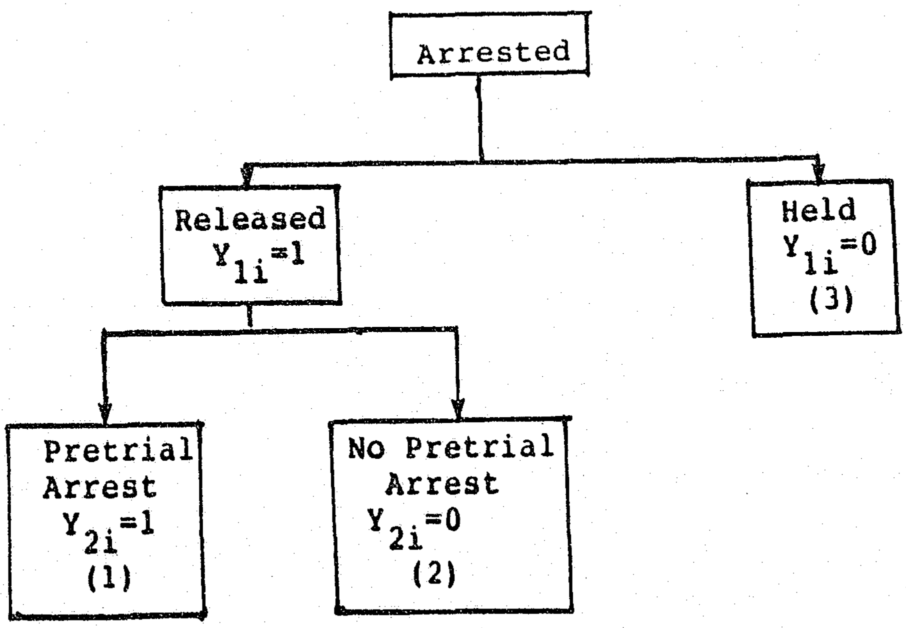
## VI. SIMPLE ILLUSTRATION OF SELECTIVITY BIAS IN A TWO-STAGE MODEL

The selectivity bias problem in pretrial release can be illustrated by setting up a simple two-stage release system. The estimation results obtained using single equation estimation techniques may be compared with those from a bivariate probit estimator capable of correcting for selectivity bias arising due to partial observability. Differences in the results illustrate the potential for incorrect inferences when selectivity problems are present.

The two-stage system selected for analysis is illustrated in Figure VI-1. The first stage is a release decision in which some accused are released, on bail or recognizance, and others are detained, in this case unable to post bond. The second stage decision, pretrial arrest, which is used as an indicator of pretrial crime, is only observed in cases where release is obtained.

FIGURE VI-1

SIMPLE TWO-STAGE RELEASE AND PRETRIAL ARREST PROCESS



Note that all arrested persons were judged to be held unless the data record contained positive evidence of release. Clearly, some persons were held for a significant period and eventually obtained release without this being recorded in the data. They are treated as held, outcome (3).

The relationships undering the flows in the pretrial arrest process shown in Figure VI-1 are similar to those shown in general form in equations (IV-1) discussed earlier and repeated below:

$$\text{(IV-1)} \quad Y_{1i}^* = G_1 + Z_{1i}g_1 + e_{1i}$$
$$Y_{2i}^* = G_2 + Z_{2i}g_2 + e_{2i}$$

where we observe $Y_{1i}$=1 if the accused is released and equal to 0 otherwise and $Y_{2i}$=1 if the accused has a pretrial arrest and 0 if no pretrial arrest occurs. This is a case of partial observability because pretrial arrest subsequent to release is not observed for cases where $Y_{1i}$=0. We expect that the system works so that persons with greater propensity for pretrial crime, i.e., persons with large $Y_{2i}^*$ and hence large expected $e_{2i}$ and $Y_{2i}$ more likely equal to 1, are also more likely to have small $Y_{1i}^*$, i.e., be less likely to secure release and hence have lower expected $e_{1i}$. Thus, we expect the correlation between $e_{1i}$ and $e_{2i}$ to be negative. This has important implications for the nature of selectivity bias, particularly affecting the estimate of the constant term $G_2$, in simple probit or ordinary least squares (OLS) estimates of the $Y_2$ equation.

The total size of the person-based sample for the first half of 1981 was 4,253 of which 2,311 cases were selected randomly, by the last digit of the police identification number, for immediate econometric analysis and 1,942 cases were kept as a holdout sample. Of the 2,311 cases, 487 were held and 1,824 were released. Exactly one-third of those released experienced pretrial rearrest, or there were 608 cases of rearrest out of 1,824 released. Appendix C provides a thorough documentation of the original data set.

Table VI-1 displays basic descriptive statistics for the entire sample of 2,311 arrest cases and for the 1,824 cases who were released, either on bail or recognizance, with or without conditions. Comparison of the averages for those released with those for all arrested, indicates, as expected, the differences in criminal history that presumably select some individuals for release. Measured in terms of number of prior convictions (Excon), fraction with pending cases (Pndcase), or fraction currently on parole (Parole), those individuals released have lower rates of past involvement with the justice system than the overall sample. A slightly higher percentage of those released were employed at time of arrest. There are no significant age, race, or gender differences between those released and the general sample. Overall, these results indicate that the severity of past criminal record is used to screen accused persons for release by judges or magistrates. Demographic factors are not important. As anticipated, this raises the potential for selectivity bias in single-equation models of release because those released differ systematically in terms of potential for pretrial arrest from those detained.

TABLE VI-1

DESCRIPTIVE STATISTICS FOR ARRESTED PERSONS RELEASED VS. HELD

| VARIABLE NAME | RELEASED | ENTIRE SAMPLE |
|---|---|---|
| PTAREST | 0.333 | 0.263 |
| RELEASE | 1.0 | 0.789 |
| AGE | 31.7 YEARS | 32.6 YEARS |
| EXCON | 2.05 CONVICTIONS | 2.42 CONVICTIONS |
| PNDCASE | 0.28 | 0.33 |
| PAROLE | 0.15 | 0.19 |
| PROBATN | 0.86 | 0.84 |
| MALE | 0.90 | 0.90 |
| BLACK | 0.95 | 0.95 |
| DRUGS | 0.26 | 0.24 |
| EMPLOYD | 0.64 | 0.61 |
| NUMBER OF OBS. | 1,824 | 2,311 |

The single-equation approach to estimating pretrial arrest equations designed to determine the factors associated with differential propensity to commit pretrial crime is to take the 1,824 observations of released persons, who had an opportunity through release to be rearrested, and estimate a pretrial rearrest equation for them. Such an equation could be estimated using ordinary least squares (OLS), or single-equation probit techniques. Examples of such estimating equations are shown in the first two columns of Table VI-2. The estimated coefficients may appear to be substantially different in magnitude but this is, in part, due to the difference in estimation technique. A good approximation is to take probit coefficients and multiply by 0.4 to get an estimate of what the equivalent OLS coefficient would be, except for the constant term where one must multiply by 0.4 and add 0.5 to obtain the equivalent OLS value. Once these adjustments are made to the probit coefficients in the second column of Table VI-2, they are not very different than the OLS estimates.

The multivariate approach to the problem, involving a bivariate probit estimate in this case of two decisions, was estimated using the bivariate probit estimator reported in Meng and Schmidt [1985] which is, in turn, based on Porier [1981]. The bivariate probit estimation results are reported in the third column of Table VI-2.

Note particularly that the estimated constant term of the bivariate probit is far larger than that of the simple probit, 1.06 vs. -0.47. This suggests that the bivariate probit will produce estimates of the expected probability of subsequent

TABLE VI-2

OLS, PROBIT, AND BIVARIATE PROBIT ESTIMATES OF PRETRIAL ARREST

Second Equation: Pretrial Arrest Equation
Estimated Coefficients With Standard Errors In Parentheses
* Indicates Significance At The 10% Level

| Independent Variable | OLS | Probit | Bivariate Probit |
|---|---|---|---|
| Constant | 0.317* | -0.476* | 1.065* |
| | (0.041) | (0.146) | (0.294) |
| Age | -0.15* | -0.629* | -0.070* |
| | (0.049) | (0.184) | (0.012) |
| $Age^2$ | 0.353* | 0.0006 | 0.00067* |
| | (0.123) | (0.0038) | (0.00013) |
| Excon | 0.043* | 0.127* | 0.133* |
| | (0.005) | (0.015) | (0.026) |
| Pendcase | 0.118* | 0.328* | 0.307* |
| | (0.019) | (0.056) | (0.065) |
| Male | -0.030 | -0.094 | -0.120 |
| | (0.037) | (0.112) | (0.113) |
| Employd | -0.037 | -0.116* | -0.122 |
| | (0.028) | (0.067) | (0.068) |
| Confid | -0.038 | -0.119 | -0.105 |
| | (0.051) | (0.070) | (0.156) |
| Violent | -0.045 | -0.125 | -0.183 |
| | (0.047) | (0.137) | (0.131) |
| Drugs | -0.029 | -0.089 | -0.084 |
| | (0.036) | (0.106) | (0.115) |
| Larceny | 0.030 | 0.074 | 0.075 |
| | (0.039) | (0.116) | (0.121) |
| Prosty | -0.030 | -1.008* | -1.071* |
| | (0.076) | (0.252) | (0.251) |
| Weapons | -0.109* | -0.354* | -0.325* |
| | (0.053) | (0.167) | (0.169) |
| Posscrim | -0.113* | -0.336 | -0.899* |
| | (0.061) | (0.184) | (0.184) |
| Miscrim | -0.121 | -0.385* | -0.352* |
| | (0.039) | (0.117) | (0.120) |
| Number of Obs. | 1,844 | 1,844 | 2,311 |
| $r_{12}$ (correlation between $e_{1i}$ and $e_{2i}$) | | | 0.0095 |
| | | | (0.029) |
| Predicted Average Probability of Rearrest For Holdout Sample | 0.081 | 0.21 | 0.35 |
| Number Of Rearrests Predicted For Holdout Sample (Pr.0.5 Out Of 1942 Cases) | 63 | 151 | 338 |

pretrial arrest that are far larger than those obtained from the simple probit. This is, of course, precisely the result that we would expect based on theory. In the case of the system analyzed here, the best risks should be among those released. As noted in the discussion of equations (1') above, $Y_{1i}$ is the release decision and equals 1 if release is secured. We expect that $e_{1i}$ and $e_{2i}$, the error terms of the release and rearrest equations, respectively, are negatively correlated, $r_{12}<0$. Whatever unobserved factors cause an accused to be released in decision 1 should be associated with low levels of pretrial arrest and hence with low values of the error in the second equation.

The arguments developed above imply that the estimated constant term, $G_2$, in a single-equation model, such as the OLS or simple probit, should be biased downward. Indeed, this apparent bias is observed by looking at the constant terms in Table VI-2 where OLS has an estimated constant of $0.32<0.50$ and simple probit has a negative constant term, while bivariate probit has a positive estimated constant. The estimate of $r_{12}$ in Table VI-2 is 0.0095; and given the large standard error of 0.029, this is not significantly different than zero. Previous work with bivariate probit estimators has had problems with estimates of the correlation between the error terms. Given that these are correlations between unobservable variables, it is not surprising that precise results are difficult to obtain. Unlike previous studies, $r_{12}$ in this analysis did not tend to leave the -1 to +1 interval where it logically should be found.

Of course, there are other differences in the estimated coefficients of Table VI-2 beyond those in the constant terms. It seems clear that sequential selectivity is strong enough in this two-stage system to have a substantial effect on estimates of the determinants of pretrial arrest. One way to summarize these differences is to predict pretrial arrest using the holdout sample of 1,942 cases.

The results of this effort are shown at the bottom of Table VI-2. First, the average probability of pretrial arrest was predicted using the holdout sample. This can be thought of as a forecast of expected rearrest if all the arrested persons were released. Note that the average probability predicted by the bivariate probit, 0.35, is far larger than the 0.21 from simple probit or 0.081 from OLS. Given that the average probability for those released was 0.33 and that those not released should have been even more likely to commit pretrial crime, estimates below 0.30 seem unrealistically low, as one might expect given the direction of the downward bias due to sample selection. Second, predictions of the number of pretrial arrests were made with an estimated probability of 0.5 or greater used to select those expected to commit crime. As expected, the predicted number of pretrial arrests was highest with the bivariate probit. The single-equation estimates certainly seem far too low and, indeed, are substantially below the actual number of 420 pretrial arrests observed for the holdout sample. Note that the 420 number was reached without allowing for the inability of those held to experience pretrial rearrest (except in extraordinary cases of arrest for an offense committed while in jail.)

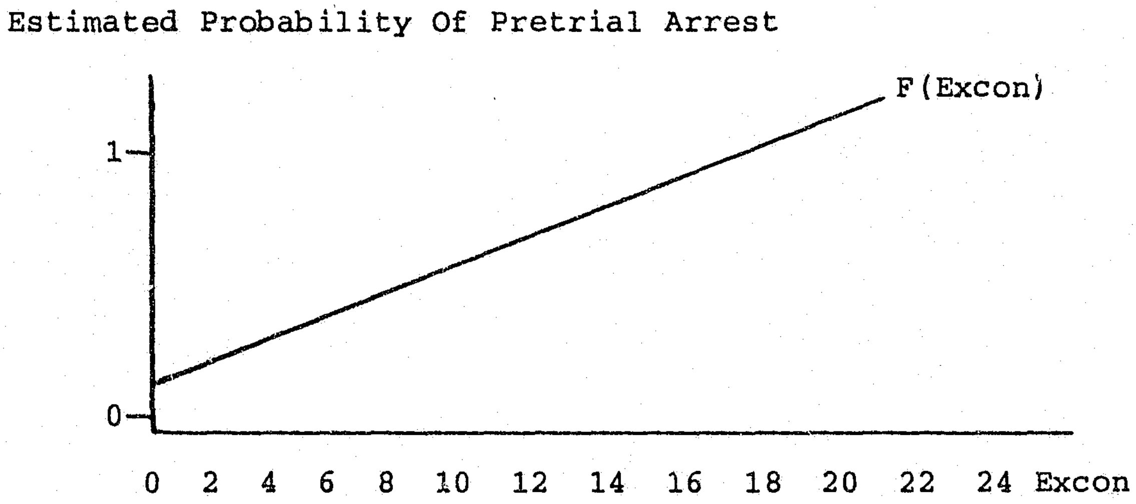
Overall, the results in Table VI-2 illustrate the danger of relying on single-equation models which are estimated using samples subject to sequential selection bias for making inferences about the determinants of misconduct in pretrial release situations. Clearly, policy based on the estimation results reported in the first two columns of Table VI-2 would be likely to produce outcomes which were not desired or anticipated. The selection bias tends to be systematic, particularly in its effect on the constant term, with the incomplete information likely to concern the behavior of high-risk cases which are not released. The estimates obtained from simple approaches tend to be systematically wrong - i.e., the resulting errors are not random. Conclusions based on results with such systematic bias are not likely to be helpful to the operation of pretrial release processes.

The estimated coefficients obtained using ordinary least squares (OLS) techniques have a more straightforward interpretation than their counterparts from either simple probit or bivariate probit estimators. The OLS estimating equation is essentially a probability "score" with each variable having an additive and independent influence on the estimated probability of pretrial crime. It is, therefore, possible to plot a simple partial relationship between any of the independent variables and predicted pretrial crime. For all variables except age, this relationship is linear and is graphed as a straight line in Figure VI-2. Such a linear relationship is illustrated for the particular case of the variable "Excon," which is the number of prior criminal convictions, in Figure VI-2.

Figure VI-2

Illustration Of OLS Estimates Of Effect of Excon

On Pretrial Arrest




The effect of Excon on predicted pretrial arrest, read as the estimated coefficient of Excon in Table VI-2, was a constant 0.043, or an increase of 4.3 percentage points for each prior conviction. For a variable like age which enters with linear and squared terms, the OLS results in Table VI-2 show that the sign on age is negative and the sign on age-squared is positive, so that the comparable function for age, F(Age), would have a "U" shape, falling at a decreasing rate over the relevant range of ages. Of course, each of these functions gives a partial relationship between the independent variable and the predicted probability of pretrial arrest.

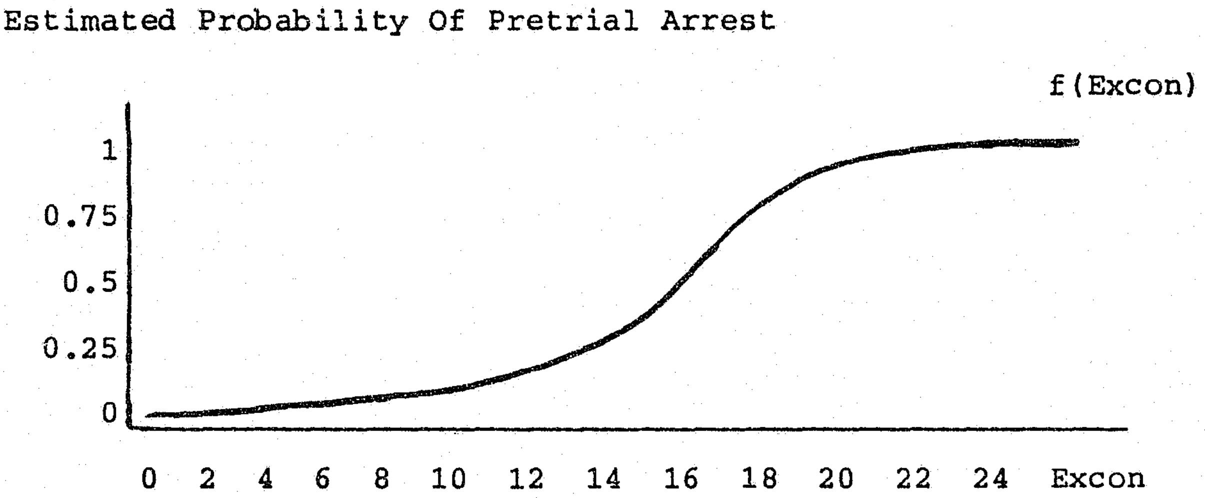
The overall estimate will depend on the "score" that is calculated by computing the arithmetic sum of the partial effects. While all this may seem to recommend the OLS approach as yielding simple, intuitive insights into the causes of pretrial arrest, there is a problem with the OLS estimates because the predicted probability may be either less than zero or greater than one, either case making no sense. Predicted probabilities obtained when scores are computed using OLS estimates do not necesarily lie on the zero - one interval. This could create problems if the scores were used directly to generate predictions for use by decision makers. Note that this problem arises naturally as a logical consequence of the way in which the OLS estimator requires that the underlying model be specified.

In contrast, the probit estimator computes the estimated probability of pretrial arrest using the cumulative normal distribution function. This makes the effects of a particular independent variable more difficult to estimate but guarantees that the predicted probabilities must lie on the zero - one interval. If the predicted probability varies with Excon according to f(Excon), then predicted probability is represented by the relation shown in Figure VI-3.

The cumulative normal function used as the basis for probit has a characteristic nonlinear "S" shape. For very low probability values, the function is very flat as shown in Figure VI-3 and changes in Excon have little effect on the predicted probability. Put another way, if other characteristics indicate that predicted pretrial arrest is unlikely, then additional prior convictions will not have an important effect on pretrial arrest. However, in an intermediate range, where other factors indicate that the predicted probability is about 0.5, the curve in Figure VI-3 is quite steep and small changes in Excon will have important effects on the predicted probability. The

Figure VI-3

Illustration Of Probit Estimates Of Effect Of Excon On Pretrial Arrest




discussion above suggested that OLS and probit coefficients could be compared by multiplying the probit coefficients by 0.4, except for the constant term where comparison is made by multiplying the constant by 0.4 and adding 0.5. The nonlinearity inherent in the probit estimate of predicted probability means that one must know the full range of characteristics of the individual in order to evaluate the marginal effect of a particular variable. For example, if the estimated probability based on the other variables were 0.25 overall, then the probit estimates suggest that a unit increase in Excon results in an increase in the predicted probability of about 0.041. However, if the estimated probability based on other variables were about 0.5, the marginal effect of Excon would rise to 0.051, a 25% increase over the probability effect at 0.25. It is important to consider this nonlinearity in the probit-predicted probabilities when interpreting the probit coefficients and comparing them to the OLS estimates.

# VII. SELECTIVITY BIAS IN THREE-STAGE MODELS OF PRETRIAL ARREST

Bivariate probit restricts our ability to estimate relationships in systems with sequential selectivity, such as pretrial release, to cases where there are two decision points. Lee's [1984] proposed method promises to allow unbiased estimation of the parameters of choice models involving several stages of selectivity. In order to test the feasibility of an operational version of Lee's approach, a fortran computer program to obtain the maximum likelihood estimates was developed. This involved precise specification of the likelihood function for the multivariate probit, differentiation of the likelihood function, and implementation of the analytical results in a fortran computer program. The algorithm used to obtain the maximum likelihood estimates is described in Berndt [1974]. The evaluation of single and double integrals was accomplished with the IMSL subroutines DCADRE and MDBNOR. The inverse normal function was computed with the IMSL subroutine MSNRIS.

The resulting software was tested using data artificially generated from a zero mean, unit variance, trivariate normal distribution with cross-equation correlation coefficients of 0.25. Even with sample sizes as small as 300, the computer program was found to produce reliable parameter estimates, although no formal monte carlo study was undertaken. The only disappointment was the failure to produce statistically significant cross-equation correlation coefficients. Although all the estimates of the correlation coefficients were close to the true value of 0.25, the largest t-statistic obtained was 1.0.

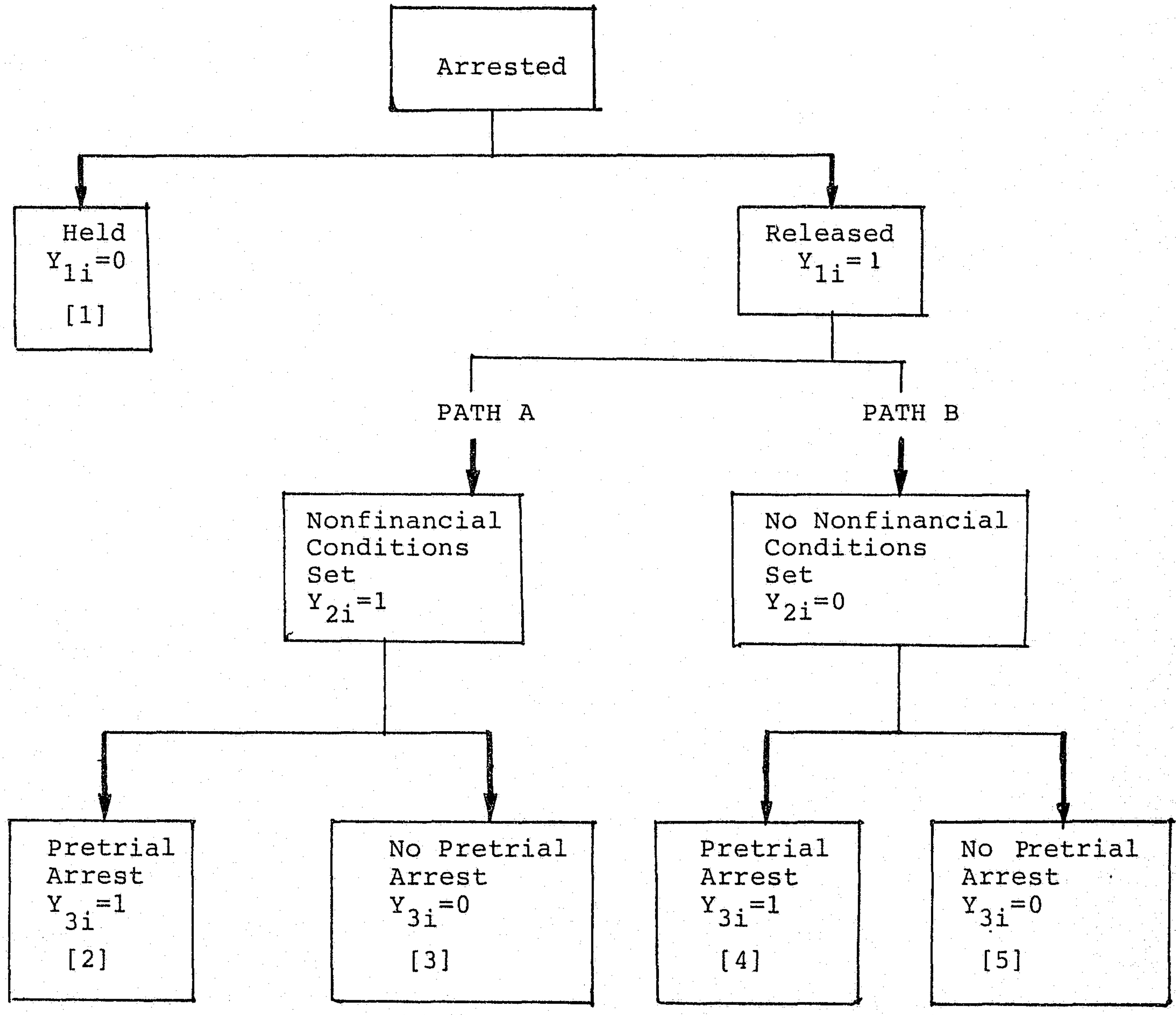
Researchers using bivariate probit estimators have also reported difficulty with estimates of cross-equation correlations, both with significance levels and with values outside the -1 to +1 range of feasibility. This second problem seldom occurred with the trivariate probit estimates of the pretrial arrest which were examined in the course of this project.

It is possible to implement several three-stage models of pretrial release, given the structure of the present system. In the remainder of this section, two interesting models that deal with the controversial role of money bail vs. nonfinancial release conditions in inhibiting pretrial crime are tested and the results presented. Generally, each experimental model produced results, particularly using multivariate vs. monovariate probit, which indicated that sequential selectivity was having a significant effect on the results of the simpler models.

The first three-stage model concentrates on the setting of release conditions and the eventual observation of pretrial arrest. This model is described in Figure VII-1 which shows that there are five possible outcomes for the accused. Partial observability arises because we cannot observe the pretrial arrest behavior of the group of arrested persons who are held and because we do not observe pretrial arrest under nonfinancial conditions for all accused but rather only for the group receiving release on nonfinancial conditions.

FIGURE VII-1

THREE-STAGE MODEL OF CONDITIONS IN PRETRIAL RELEASE AND ARREST



Note that there are really two complete trivariate processes in Figure VII-1. One consists of the system where nonfinancial conditions are set, $Y_{1i}=Y_{2i}=1$, and pretrial arrest behavior in outcomes (2) and (3) is observed, $Y_{3i}=0,1$. The other is based on pretrial arrest of persons released with no nonfinancial conditions, outcomes (4) and (5) $Y_{1i}=1$ and $Y_{2i}=0$, in Figure VII-1. In subsequent discussion, these will be termed path A and path B respectively. Estimates performed on path A indicate the determinants of pretrial arrest among accused individuals who were released with nonfinancial conditions set.

In contrast, estimates on path B allow the prediction of pretrial arrest associated with individuals released with no nonfinancial conditions set, i.e. cases in which bail was set or cases with outright, unconditional release. Note that the setting of bail is usually viewed as a way of ensuring appearance for trial, not as a way of reducing pretrial crime. Therefore, it is certainly possible that setting financial conditions is not an important determinant of pretrial arrest.

The system in Figure VII-1 may be illustrated using equations (3) shown below. The actual outcomes in Figure VII-1 are structured so that, if the accused is released, $Y_{1i}=1$; and $Y_{1i}=0$, if the accused is held.

(VII-1)
$$Y_{1i}^* = G_1 + Z_{1i}g_1 + e_{1i}$$
$$Y_{2i}^* = G_2 + Z_{2i}g_2 + e_{2i}$$
$$Y_{3i}^* = G_3 + Z_{3i}g_3 + e_{3i}$$

The outcome in which nonfinancial conditions are set is realized if $Y_{2i}=1$ and release without such conditions if $Y_{2i}=0$. Finally, $Y_{3i}=1$ for the cases in which pretrial arrest occurs. This system has two levels of selectivity and three possibilities for correlation between the error terms. We suspect that individuals who are held are the worst risks and, indeed, judges may anticipate future pretrial crime problems in making release decisions. Thus, we anticipate that the correlation between $e_{1i}$ and $e_{3i}$, $r_{13}$, will be negative: any accused with a large positive value of $e_{3i}$ will tend to be perceived as a poor risk for release and hence likely to be held.

Put another way, an omitted variable which enters $e_{3i}$ so that it varies directly with the implicit probability of pretrial crime is likely to vary inversely with the implicit probability of release in the first equation. If the accused with the highest risks for pretrial crime are selected out of the sample because they are held, then single-equation estimates of pretrial crime determinants on either path A or B will tend to understate the likely amount of pretrial crime that would occur if all accused were released. This analysis suggesting that $G_{3i}$ might be biased downward is too simplistic because it ignores $r_{23}$ and $r_{12}$ which also influence the selectivity bias in estimates of $G_{3i}$. If those released with nonfinancial conditions are generally better risks, then $r_{12}$ will be positive and $r_{23}$ will be negative. The net effect of these potential sources of bias on the magnitude of $G_3$ is an empirical question.

The estimation results for this model of nonfinancial conditions and pretrial arrest for both path A (nonfinancial conditions set) and path B (no nonfinancial conditions set) are shown in Table VII-1. The results of three estimation techniques are shown: ordinary least squares (OLS), single-equation

TABLE VII-1
OLS, PROBIT, AND TRIVARIATE PROBIT ESTIMATES OF PRETRIAL ARREST AND EFFECTS OF SETTING CONDITIONS FOR RELEASE

Third Equation: Pretrial Arrest Equation Paths A And B
Estimated Coefficients With Standard Errors In Parentheses
* Indicates Significance At The 10% Level

| Independent | OLS | | Probit | | Trivariate Probit | |
|---|---|---|---|---|---|---|
| Variable | Path A | Path B | Path A | Path B | Path A | Path B |
| Constant | 0.861* | 0.801* | 1.153* | 0.980* | 1.284* | 0.587 |
| | (0.119) | (0.174) | (0.359) | (0.528) | (0.544) | (0.702) |
| Age | -0.024* | -0.023* | -0.073* | -0.067* | -0.072* | -0.068* |
| | (0.005) | (0.007) | (0.015) | (0.021) | (0.016) | (0.023) |
| $Age_2$ | 0.00022* | 0.00021* | 0.00067* | 0.00060* | 0.00066* | 0.00057* |
| | (0.00005) | (0.00007) | (0.00016) | (0.00022) | (0.00018) | (0.00022) |
| Excon | 0.044* | 0.050* | 0.131* | 0.149* | 0.128* | 0.135* |
| | (0.006) | (0.008) | (0.020) | (0.025) | (0.034) | (0.036) |
| Pendcase | 0.119* | 0.131* | 0.329* | 0.367* | -0.339* | 0.332* |
| | (0.025) | (0.032) | (0.079) | (0.095) | (0.093) | (0.098) |
| Male | 0.0039 | 0.051 | -0.006 | 0.127 | -0.108* | -0.268 |
| | (0.048) | (0.063) | (0.142) | (0.197) | (0.138) | (0.191) |
| Emplyd | -0.061* | -0.047 | -0.183* | -0.150 | -0.184* | -0.011 |
| | (0.029) | (0.037) | (0.085) | (0.113) | (0.085) | (0.107) |
| Confid | -0.050 | -0.028 | -0.143 | -0.090 | -0.136 | 0.089* |
| | (0.065) | (0.082) | (0.189) | (0.242) | (0.206) | (0.238) |
| Violent | -0.099* | -0.113 | -0.278 | -0.317 | -0.266 | 0.007 |
| | (0.054) | (0.074) | (0.159) | (0.223) | (0.183) | (0.278) |
| Drugs | -0.019 | -0.025 | -0.057 | -0.072 | -0.068 | -0.104 |
| | (0.046) | (0.059) | (0.132) | (0.171) | (0.157) | (0.207) |
| Larceny | -0.004 | -0.035 | -0.022 | -0.111 | -0.026 | 0.249 |
| | (0.054) | (0.065) | (0.147) | (0.190) | (0.157) | (0.231) |
| Prosty | -0.358* | -0.323* | -1.166* | -1.045* | -1.166* | -0.771 |
| | (0.097) | (0.127) | (0.318) | (0.419) | (0.314) | (0.408) |
| Weapons | -0.152* | -0.214* | -0.464* | -0.675* | -0.472* | -0.081 |
| | (0.077) | (0.092) | (0.236) | (0.295) | (0.256) | (0.274) |
| Posscrim | -0.153* | -0.248* | -0.435* | -0.756* | -0.450* | -0.106 |
| | (0.083) | (0.106) | (0.226) | (0.336) | (0.250) | (0.290) |
| Miscrime | -0.139* | -0.177* | -0.437* | -0.562* | -0.435* | -0.125 |
| | (0.048) | (0.061) | (0.142) | (0.187) | (0.147) | (0.197) |
| Cross Equation Correlation Coefficients | | | | | | |
| $r_{12}$ | | | | | 0.202 | -0.522* |
| | | | | | (0.254) | (0.233) |
| $r_{13}$ | | | | | -0.004 | -0.342 |
| | | | | | (0.553) | (0.630) |
| $r_{23}$ | | | | | -0.153 | 0.450 |
| | | | | | (0.502) | (0.498) |
| Predicted Average Pretrial Arrest Rate For Holdout Sample | | | | | | |
| | 0.315 | 0.248 | 0.373 | 0.322 | 0.419 | 0.232 |
| Number of Pretrial Arrests Predicted For Holdout Sample | | | | | | |
| | 213 | 183 | 369 | 221 | 585 | 112 |
| (Pr > 0.5, Out Of 1942 Cases) | | | | | | |

probit, and trivariate probit. The two single-equation techniques are estimated for very different samples, path A uses the 1138 cases where nonfinancial conditions were set and path B estimates are based on 686 releases where no nonfinancial conditions were imposed. Of course, neither of these single-equation approaches considers any of the 487 cases in which the accused is held, while the trivariate approach uses the entire sample of 2,311 arrested persons.

The pretrial arrest equation estimates appear to be similar. It is important to note that probit coefficient estimates should be multiplied by 0.4 (for the constant term, multiply by 0.4 and add 0.5) to make them comparable to OLS-estimated coefficients, which can be interpreted as incremental contributions to the probability of pretrial arrest. Once the probit estimates from the single or trivariate technique have been adjusted, they can be compared directly to OLS and their implication for differences in the expected probability of pretrial arrest due to differences in characteristics of the accused appreciated. Thus, if a probit coefficient were 0.2, then multipling by 0.4 gives 0.08 and suggests that a unit increase in the variable associated with that coefficient will increase the probability of pretrial arrest by 0.08.

As might be expected, pretrial arrest probability decreases (at a decreasing rate) with age and is lower for those who are currently employed. The probability of pretrial arrest increases with the number of prior convictions and the number of pending cases. Also, certain types of crime appear to be more regularly related to pretrial arrest than other types. Prostitution, weapons violations, possession of the instruments of crime, and miscellaneous crimes including auto theft, shoplifting, and possessing stolen property are all categories of charge at arrest fairly consistently associated with lower probability of pretrial arrest.

Clearly, there are differences among the estimates of these pretrial arrest probability equations based on estimation technique. But there are also differences between results for path A and path B. It is difficult to summarize these differences but one may consider, for example, the differences in constant terms, which tend to reflect the effects of selectivity bias most directly. The constant terms for estimates of path A are larger than those for path B, but the differences between A and B constant terms are small for the OLS and probit estimators and large for the trivariate probit.

Based on the arguments above, we had anticipated that the $G_3$ constant terms for the single equation estimates of path A would be biased downward while the single equation estimates of the constant term for path B might be biased upward, although this conclusion was rather tentative. The final results in Table VII-1 do show this pattern with trivariate probit estimates of the path A constant term being above those obtained with the path A single-equation approaches. In contrast, trivariate probit estimates of the constant term for path B are lower than constant terms estimated for the single-equation estimators.

As anticipated in the above discussion, $r_{12}$ for path A is positive. Omitted factors entering $e_{1i}$ which judges perceive to be indicators that the accused is a good risk are positively correlated with the omitted factors causing setting of nonfinancial conditions. By the same line of argument, $r_{12}$ is negative for path B. This is the only cross-equation correlation found to be statistically significant. As noted above, the general lack of significance in the estimated cross-equation correlation coefficients in this report may be explained by limitations on the estimation technique because informal monte carlo experiments indicated that the trivariate probit produced unbiased estimates of the $r_{ij}$'s but that the standard errors for these estimates were large.

Rather than comparing individual coefficient estimates or constant terms, the variation in the estimation results presented in Table VII-1 may be appreciated and evaluated by computing estimated average probabilities of pretrial arrest using a holdout sample of arrest cases. The holdout sample of 1,942 arrests was selected randomly from an initial group of 4,253 arrests (the other 2,311 were used to construct the estimates), as described in Chapter V, above.

The average estimated probability of pretrial arrest is computed for this holdout sample and recorded at the bottom of Table VII-1. The differences in expected pretrial arrest between path A and B are small for the OLS, 0.32 for path A and 0.25 for path B. The difference is almost identical for probit estimates, 0.37 for path A vs. 0.32 for path B. These results may reflect the expected downward bias in estimates from path A and upward bias in path B. However, the trivariate probit shows very different rates of pretrial arrest, 0.42 for path A with nonfinancial conditions set vs. 0.23 for path B with no nonfinancial conditions set (release on bail or unconditional release).

These estimated rates of pretrial arrest from the trivariate probit are the unconditional expectation of the rate of arrest if all the 1,942 arrested persons were forced down a particular release path taking into account the selectivity in release path in the original data set. The trivariate probit results suggest that releasing all accused persons on nonfinancial conditions would result in substantial rates of pretrial arrest compared to path B in which all the accused are released on bail or released unconditionally.

Single-equation models do not show significant differences in the pretrial arrest rate due to differences in release procedure. Indeed, they do not generate estimates of rates of pretrial arrest that are significantly above the 0.33 rate which was observed in the current data set for those released on conditions. Of course, those held without release should have the highest arrest rate, and it is the downward bias in the single-equation estimates of pretrial arrest that accounts for the low predicted rearrest rate when the entire sample of 1,942 is forced through path A.

A final way to evaluate the differences in predictions of pretrial arrest across paths and estimation techniques is to use the holdout sample to predict the expected number of pretrial arrests from the total of 1,942 arrested persons. If the estimated probability of pretrial arrest for a case in the holdout sample is greater than 0.5, then this is counted as a predicted pretrial arrest. The results at the bottom of Table

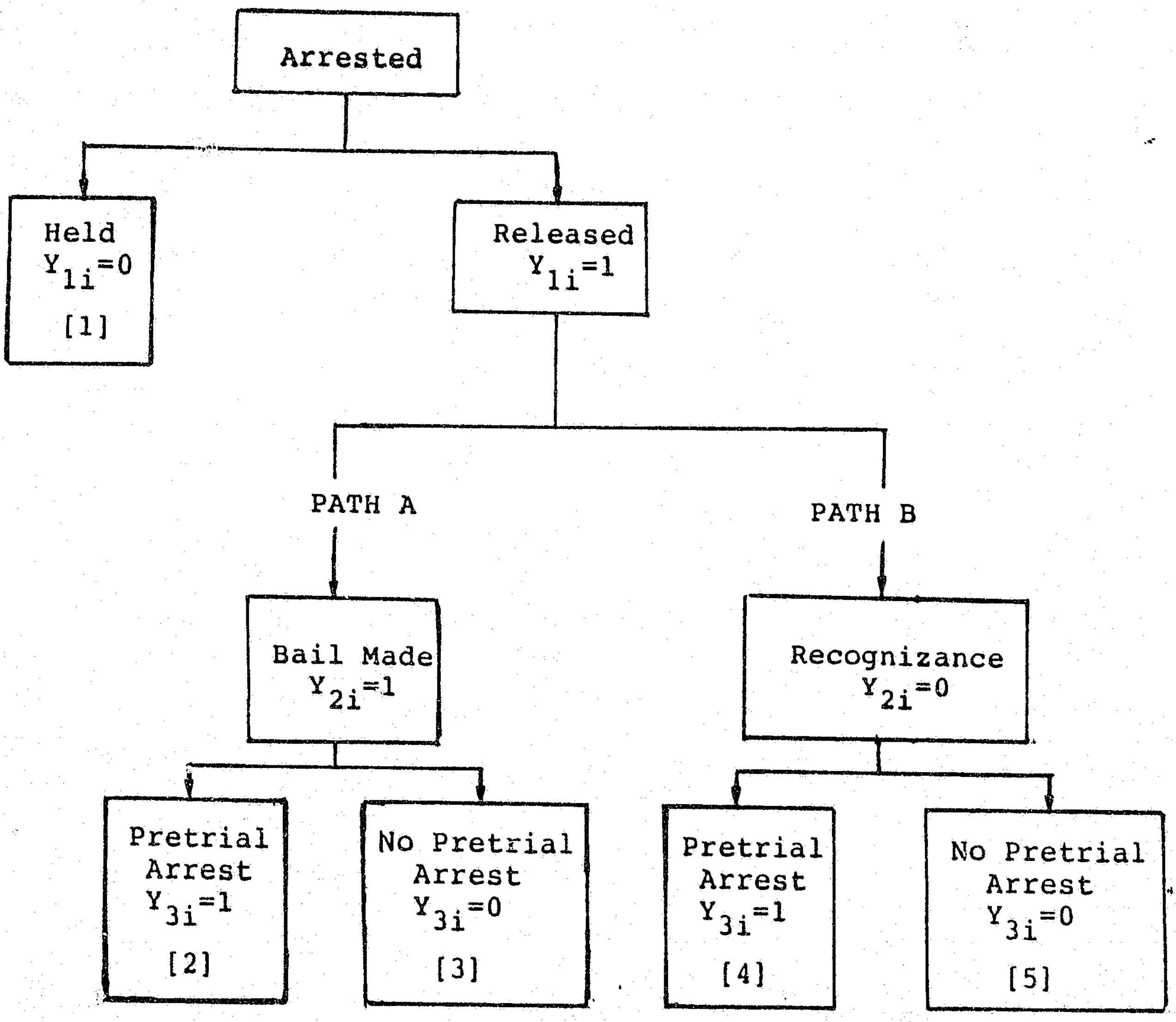
VII-1 show very dramatic differences in total pretrial arrests associated with path A (585 arrests) vs. path B (112 arrests) for the trivariate probit model. These differences are greatly reduced in the single-equation results, where OLS predicts 213 and 183 and probit predicts 369 and 221 pretrial arrests for paths A and B, respectively.

If policy recommendations concerning the effects of release conditions on pretrial arrest were based on the single-equation model results, significant errors might result. As with the effects of selectivity bias found in the binary probit model above, this three-stage model shows that partial observability and sequential selection in pretrial release systems can have a substantial effect on the results of statistical estimation.

The second three-stage model of pretrial arrest to be considered is particularly directed to the question of release on bail vs. release on recognizance. The structure of the model is outlined in Figure VII-2. As with the earlier model, there are two paths which lead to observable pretrial arrest behavior: path A with $Y_{1i}=Y_{2i}=1$ leads to outcomes (2) and (3), and path B with $Y_{1i}=1$ and $Y_{2i}=0$ leads to outcomes (4) and (5). Single equation methods can be used to estimate pretrial arrest equations for these two paths using the 286 cases which followed path A, bail made, or the 1538 cases of release on recognizance, path B . Alternatively, trivariate estimates of the pretrial crime equation appropriate for each path can be estimated using the entire sample of 2,311 which includes those held and not making bail.

FIGURE VII-2

THREE-STAGE MODEL OF BAIL SETTING AND PRETRIAL ARREST



If judges are making accurate forecasts of risk when they make release decisions, we would expect the best risks to follow path B ($Y_{1i}=1$, $Y_{2i}=0$), intermediate risks to be in path A ($Y_{1i}=Y_{2i}=1$), and highest risks to have outcome (1). The observed rate of pretrial arrest for those in path A was 0.378 which is indeed greater than the 0.325 for those in path B. This difference reflects both selectivity which generates different samples on the two paths and also the effects of differences in treatment of those gaining release on recognizance vs. by posting bond. The basic expectation for the direction of selectivity bias as it would affect the estimated constant term is for estimates of path B to be biased downward because $r_{13}$ and $r_{23}$ should be negative. The omitted factors that cause an accused to be viewed as a good enough risk to be released in general and specifically to be released on recognizance should be negatively correlated with the decision to engage in pretrial crime. The direction of bias in single-equation estimates of path A is ambiguous.

Table VII-2 contains the results of OLS, single equation probit, and trivariate probit estimates of the model displayed in Figure VII-2. In contrast to expectations, the estimated constant terms for OLS and probit techniques are larger for path B than for path A. Also, comparing the estimated constant for path B using trivariate probit with that for simple probit, $1.110<1.273$, we find a small and non-significant decrease in the estimation result using trivariate probit. The above discussion suggested that single-equation estimates of path B should be biased downward, not that they should be too high.

TABLE VII-2
OLS, PROBIT, AND TRIVARIATE PROBIT ESTIMATES OF PRETRIAL ARREST
EFFECTS OF RELEASE ON RECOGNIZANCE VS. BAIL

Third Equation Of Model: Pretrial Arrest Equation Paths A And B
Estimated Coefficients With Standard Errors In Parentheses
* Indicates Significance At The 10% Level

| Independent | OLS | | Probit | | Trivariate Probit | |
|---|---|---|---|---|---|---|
| Variables | Path A | Path B | Path A | Path B | Path A | Path B |
| Constant | 0.515* | 0.866* | 0.069 | 1.273* | 0.263 | 1.110* |
| | (0.261) | (0.107) | (0.714) | (0.322) | (1.873) | (0.395) |
| Age | -0.008 | -0.026* | -0.023 | -0.081* | -0.021 | -0.082* |
| | (0.010) | (0.004) | (0.029) | (0.014) | (0.035) | (0.015) |
| $Age^2$ | 0.0008 | 0.0002* | 0.0002 | 0.0007* | 0.0002 | 0.0007* |
| | (.00009) | (.00004) | (0.0002) | (0.00015) | (0.0003) | (0.0002) |
| Excon | 0.028* | 0.051* | 0.079* | 0.157* | 0.158* | 0.145* |
| | (0.011) | (0.006) | (0.032) | (0.018) | (0.070) | (0.026) |
| Pendcase | 0.046 | 0.138* | 0.132 | 0.385* | -0.031 | 0.415* |
| | (0.044) | (0.022) | (0.121) | (0.067) | (0.069) | (0.160) |
| Male | -0.152 | -0.009 | -0.422 | -0.056 | -0.005 | 0.031 |
| | (0.104) | (0.040) | (0.281) | (0.125) | (0.365) | (0.123) |
| Employd | 0.043 | -0.049* | 0.118 | -0.158* | -0.354* | -0.199 |
| | (0.061) | (0.024) | (0.108) | (0.076) | (0.213) | (0.155) |
| Drugs | 0.162* | 0.070* | 0.435* | 0.213* | 0.494* | 0.219 |
| | (0.094) | (0.038) | (0.246) | (0.117) | (0.208) | (0.327) |
| NOB | 286 | 1,538 | 286 | 1,538 | 2,311 | 2,311 |
| Cross Equation Correlation Coefficients | | | | | | |
| $r_{12}$ | | | | | 0.220 | -0.211 |
| | | | | | (0.329) | (0.331) |
| $r_{13}$ | | | | | 0.292 | 0.254 |
| | | | | | (0.674) | (0.323) |
| $r_{23}$ | | | | | -0.182 | -0.291 |
| | | | | | (1.207) | (1.236) |
| Predicted Average Pretrial Arrest Rate For Holdout Sample | 0.36 | 0.24 | 0.40 | 0.31 | 0.430 | 0.390 |
| Predicted Number of Pretrial Arrests For Holdout Sample (Pr>0.5 Out Of 1,942 Cases) | 395 | 174 | 420 | 231 | 524 | 470 |

The apparent mystery is solved by looking at the predicted average pretrial arrest rates and number of pretrial arrests at the bottom of Table VII-2. Predicted pretrial arrest, whether rate or number out of the holdout sample of 1,942, is significantly higher for the single-equation estimates of path A than for path B. Apparently, differences between path A and B in the estimated coefficients of the independent variables, particularly differences in the coefficient of age, were so large and in the opposite direction of differences in constant term estimates that the net effect was to produce estimated probabilities of pretrial arrest that were much higher for path A, release on bail, than for path B, release on recognizance. This, of course, agrees with our expectation that path estimates should be biased downward.

The predictions of average pretrial arrest and numbers of arrests also show that trivariate probit results move as expected. The path B trivariate predicted rate and number of arrests is significantly larger than the single-equation results, reflecting the ability of the trivariate probit to adjust for selectivity bias. Comparing the trivariate probit results for paths A and B, there are only small differences in the average probability and predicted number of pretrial arrests in the holdout sample. This suggests, in contrast to the large differences in single-equation estimates, that the predicted differences in pretrial arrest rates if one forced the entire sample of 1,942 either through release on recognizance or through release on bail would be small. But pretrial arrest rates would be slightly lower for those released on recognizance.

The finding that nonfinancial release conditions, which are imposed on those released on recognizance far more often than on those released on bail, may lower rates of pretrial arrest slightly is not surprising. Bail is designed to promote appearance for trial, not to deter pretrial crime. Nonfinancial conditions may limit opportunities for criminal behavior.

The variables which are significant predictors of differences in pretrial arrest in Table VII-2 generally have the expected effect. Prior involvement in crime and the criminal justice system, as reflected in number of prior convictions (EXCON) or number of pending cases (PENDCASE), increases expected arrest probability. Having current employment lowers that probability as does increasing age. Differences in pretrial arrest probability by type of crime were small for this model, and all of the crime type variables from Table VII-1 were dropped due to nonsignificance except the dummy variable indicating a drug charge.

As with the previous three-stage model, the results presented here indicate that there is substantial potential for simple single-equation analysis of pretrial release systems to produce seriously biased results. There has been much debate about the role of bail vs. release on recognizance, i.e., path A vs. path B. To the extent that arguments are based on either simple differences in average observed pretrial arrest rates or on the type of statistical analysis presented here as OLS or single-equation probit, this debate has been misinformed. The multivariate probit estimator developed here has the potential to produce estimates which give unconditional predictions of the differences in pretrial misconduct. These unconditional predictions can be used to predict the outcome expected if all accused persons entering into the pretrial justice system were given the same type of treatment. The unconditional predictions can be made even if the raw data used in the analysis come from a pretrial justice system that places accused persons in different treatment groups based on their personal characteristics. Single-equation approaches yield only conditional estimates which predict behavior of accused persons conditional on the choice process which selects the accused into different groups. The results above suggest that conditional results differ significantly from unconditional estimates.

# VIII. SELECTIVITY BIAS IN THREE-STAGE MODELS OF FAILURE TO APPEAR

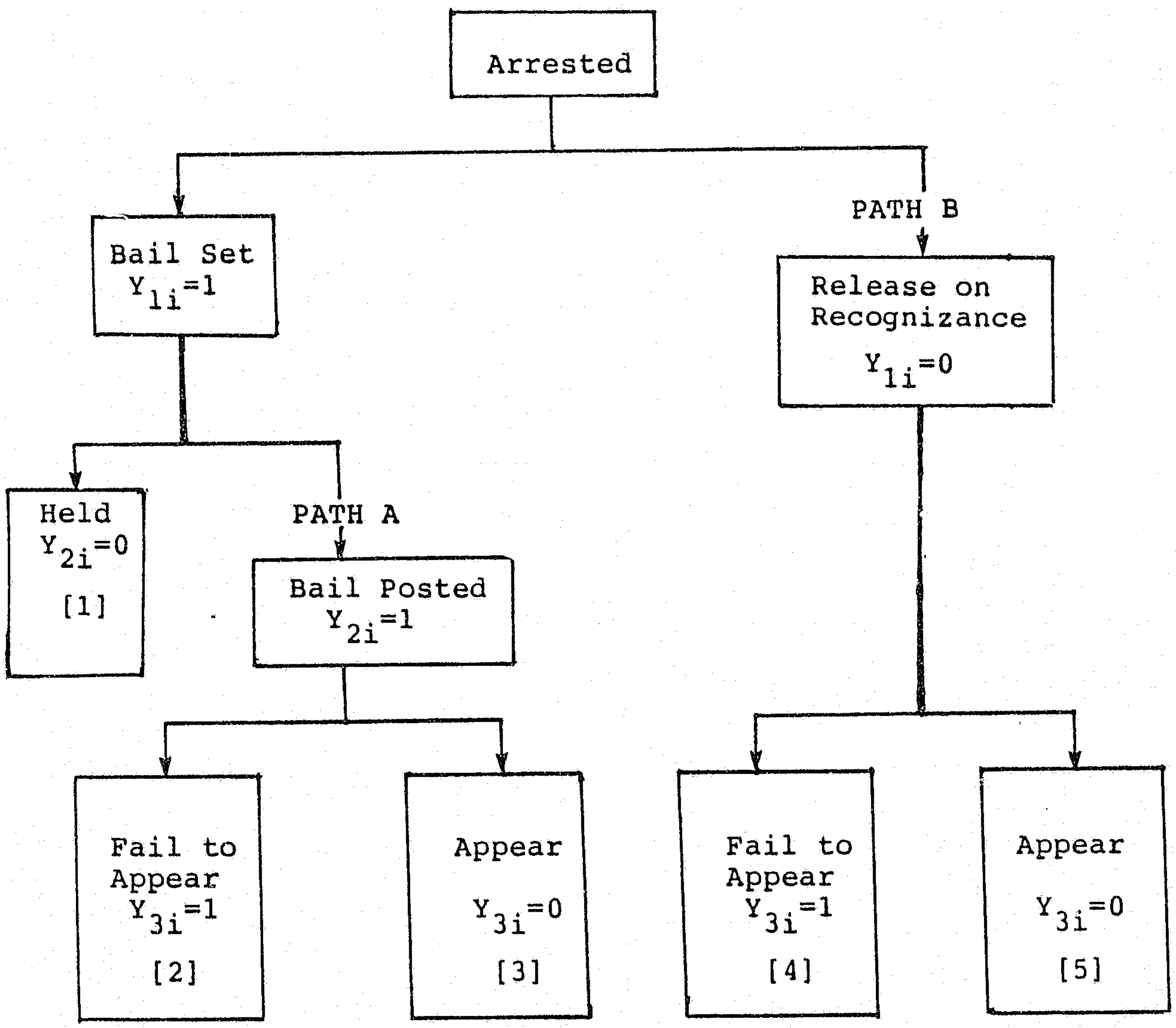
Previous sections have concentrated on the analysis and prediction of pretrial arrest and the role of selectivity bias in affecting estimates of the effects of release conditions and bail setting. Of course, a primary focus of release conditions and bail setting is the avoidance of failure to appear. This section considers specifically the classic problem of estimating the effect of bail setting on appearance. Because failure to appear is an infrequent event, this provides an excellent test of Lee's proposed method in low probability estimation.

There is substantial *a priori* information suggesting that sample selection problems may have an important influence on single-equation estimation results of failure to appear. The arguments supporting the presence of selectivity bias are apparent if one examines the flow of accused persons through the justice system as depicted in Figure VIII-1. The first level of screening is performed by the judge who attempts to isolate persons whose probability of appearance is highest and grant them release on recognizance.

In the second stage arrested persons for whom bail was set either succeed in posting bond or are held. Theory suggests several possible arguments could be made about the nature of selectivity problems which occur at this stage. First, posting bond may act to deter failure to appear, either because the accused would anticipate possible loss of funds or because the bondsman could impose costs on the accused. Such costs could range from the loss of deposits with the bondsman to the personal

FIGURE VIII-1

THREE-STAGE MODEL OF BAIL SETTING AND FAILURE TO APPEAR



efforts of the bondsman to insure that there is appearance or that initial failure to appear is remedied. Second, posting bond may indicate that the accused has significant amounts of personal resources. Such command over wealth may be associated with responsible appearance behavior. Third, posting bond usually implies that the accused is able to convince others - family, friends, or a bondsman - that appearance is likely. In effect, these other individuals act as a screening device to exclude individuals who are most likely to fail to appear. All three of these arguments suggest that, of the persons for whom bail is set, those posting bond are less likely to fail to appear.

There is one effect that could produce higher rates of failure to appear among those posting bond. Individuals fearing conviction and subsequent punishment may post bond in order to flee. The relative importance of this effect is likely to be small in comparison to the factors promoting appearance among those posting bond. The discussion will thus be based on the expectation that group [1] in Figure VIII-1 would have the highest risk of failure to appear, followed by those released on bond, path A. Those released on recognizance, path B, are expected to have the lowest risk of failure to appear.

The system shown in Figure VIII-1 may be illustrated using equations VIII-1 shown below. The outcomes in Figure VIII-1 are

$$\text{(VIII-1)} \quad \begin{aligned} Y_{1i}^* &= G_1 + Z_{1i}g_1 + e_{1i} \\ Y_{2i}^* &= G_2 + Z_{2i}g_2 + e_{2i} \\ Y_{3i}^* &= G_3 + Z_{3i}g_3 + e_{3i} \end{aligned}$$

arranged so that, if the accused is released on recognizance, $Y_{1i}=0$, and $Y_{1i}=1$ if bail is set. The outcome in which bond is not met is noted $Y_{2i}=0$, and if bond is posted $Y_{2i}=1$. This decision, of course, is only observed on path A. Finally failure to appear outcomes are indicated by outcome $Y_{3i}=1$ and appearance by $Y_{3i}=0$.

Based on arguments made above, then, we anticipate that the correlation between $e_{1i}$ and $e_{2i}$, $r_{12}$, will be negative. An accused with a large positive value of $e_{1i}$ is evaluated by the judge as being very risky, and such a person is not likely to be successful in getting release on bond. Part of the reason for this conclusion is that the bail system, if it works as intended, will succeed in denying release to those most likely to fail to appear. The same characteristics that are not observed, and hence included in $e_{1i}$, which cause the judge to deny release on recognizance should also reduce the likelihood of release on bond.

The correlation between $e_{1i}$ and $e_{3i}$, $r_{3i}$, should certainly be positive if the judge is making release decisions which anticipate greater risk based on characteristics which are not observed. Finally, the correlation between $e_{2i}$ and $e_{3i}$, $r_{23}$, is most difficult to determine because the manner in which selectivity affects the bonding system is not clear. However, the general conclusion based on arguments made above is that those with unobservable characteristics indicating greater likelihood of failure to appear are less likely to achieve release on bond. This suggests that $r_{23}$ is negative.

An additional element of the system in Figure VIII-1 is the structure of path B. This path, which consists of the release decision and appearance decision, is really a two-stage process.

This path is properly estimated by the bivariate probit techniques discussed and illustrated in Chapter VI. The correlation between $e_{1i}$ and $e_{3i}$ on this path, $r^*_{13}$, is also positive by the arguments made above.

In previous chapters, both paths B and A have been three-stage processes and their single and trivariate estimates have been presented and compared. For this failure to appear analysis, estimates of failure to appear equations for path A using trivariate probit techniques will be compared to bivariate probit estimates of failure to appear for path B.

Based on the sample selection arguments, failure to appear estimates made using single-equation models applied to data from path A should predict significantly lower rates of failure to appear than comparable single-equation models using only data from path B. This has important implications for debates over the effects of bail on failure to appear. Because they deal with a sample of accused selected for their high risk of failure to appear, bondsmen may experience high rates of non-appearance in spite of sincere attempts to promote appearance.

Thus, simple comparison of unadjusted rates of failure to appear experienced among those released on recognizance with rates for those released on bond will bias conclusions against the effectiveness of the bail system. Even more elaborate statistical analysis, including estimation of single-equation models of failure to appear, will similarly generate estimates of failure to appear conditional on use of bondsmen which are biased upward compared to estimates performed on those released on recognizance.

Simple descriptive data on the characteristics and behavior of accused persons in path A and B are compared in Table VIII-1. As anticipated, failure to appear is higher for path A, release on bond, than for path B, release on recognizance. But the margin of difference, 17.4% vs. 16.1% is certainly not large, particularly considering further evidence in Table VIII-1 that the accused released on recognizance have less serious criminal histories. Note that group [1] from Figure VIII-1, those with bail set who did not post bond, does not enter the failure to appear debate because they have no opportunity to record such violations.

The comparison of the average characteristics of the accused in path A with those in path B (see Table VIII-1) shows that the justice system in general and judges in particular are sending those accused with less serious criminal records to the group released on recognizance. For example, the average number of prior convictions for those released on bond is 2.95 vs. 2.18 prior convictions for the average person released on recognizance. This difference of about 30% in average prior convictions is particularly important because the variable based on this measure, Excon, has proved to be positively related to both pretrial arrest and to failure to appear in this and other studies.

A slightly higher percentage, 20.2% vs. 18.7%, of accused on path A were on parole when arrested. Significantly more, 63.2% vs. 51.9% of those released on recognizance were currently employed when they were arrested. Finally, the percentage of those charged with drug violations was much higher among those

TABLE VIII-1

MEAN VALUES OF VARIABLES IN PATH A AND B OF FAILURE TO APPEAR MODEL

| VARIABLE | PATH A | PATH B |
|---|---|---|
| Bondset | 100% | 0% |
| Bond Posted | 100% | 0% |
| Failure To Appear | 17.4% | 16.1% |
| Age | 31.4 years | 32.8 years |
| Excon | 2.95 | 2.18 |
| Parole | 20.2% | 18.7% |
| Male | 88.2% | 89.9% |
| Employed | 51.9% | 63.2% |
| Drug Crime | 27.5% | 21.8% |
| Family Count In Area | 2.0 | 2.1 |
| Bond Amount | $3761 | ... |
| Number of Observations | 287 | 1,344 |

released on bond. The three characteristics which are similar between paths A and B are age, percent male, and number of family members in the area. But statistical analysis performed here does not indicate that these three variables are particularly important in accounting for failure to appear. Thus, it appears that judges are selecting arrested persons for release on recognizance based on such characteristics as number of prior convictions, employment, parole status, and type of crime.

A more detailed discussion of the factors which are used in making release decisions is presented in Appendix A, where single-equation statistical estimates are reported. Of course, models of the first decision, release or bail, may be estimated by single-equation techniques without bias because there is no selection - the decision is made for all arrested persons. Given the focus of this report on statistical and econometric methods for treating selectivity bias in estimates of behavorial equations in the criminal justice system, cases in which the single-equation approach is unbiased are not given detailed attention in the body of the report. However, a review of the results in Appendix A confirms the conclusions that are apparent from Table VIII-1. The estimated probability of release on recognizance falls with such factors as number of prior convictions, prior parole status, and seriousness of charge. It falls for those employed at arrest but is not significantly influenced by the number of relatives living in the area.

The determinants of failure to appear were analyzed for arrested persons sent down either path A or path B in the pretrial release system described by Figure VIII-1. Single-equation estimation techniques, both ordinary least squares and binary probit, were used to estimate failure to appear equations involving path A, outcomes [2] or [3] in Figure VIII-1, and those involving path B, outcomes [4] or [5] in Figure VIII-1. The arguments made above suggest that use of single-equation approaches for these failure to appear equations will result in selectivity bias in the estimated coefficients. It is expected that the estimated constant term, $G_3$, for path A will tend to be biased upward because $r_{13}$ should be positive, and the estimate of $G_3$ for path B should be biased downward by the sample selection due to the positive $r_{13}$. Comparing the OLS and simple probit results with those obtained with trivariate probit allows us to evaluate the magnitude of the effects of selectivity bias due to differential treatment of accused persons.

The estimation results for single equation models, both OLS and probit, are shown in Table VIII-2. A first striking result is the generally low levels of statistical significance for the estimated coefficients. In part, this may be due to the selectivity bias which tends to eliminate heterogeneity in the subsamples taking path A vs. B. The estimated coefficients differ between the path A and path B results; however, this could be due to differences in the incentives for failure to appear that accompany release on bond as opposed to release on recognizance.

As anticipated, the estimated constant term in the OLS results is much larger for path A than for path B, indicating the counter-intuitive result that setting bail tends

TABLE VIII-2
OLS, PROBIT, AND TRIVARIATE PROBIT ESTIMATES OF DETERMINANTS OF FAILURE TO APPEAR
Third Equation Of The Model: Failure To Appear Paths A And B
Estimated Coefficients With Standard Errors In Parentheses
* Indicates Estimated Coefficient Significant At 10% Level

| Independent | OLS | | Probit | | Multivariate Probit | |
|---|---|---|---|---|---|---|
| Variables | Path A | Path B | Path A | Path B | Path A | Path B |
| Constant | 0.200<br>(0.204) | 0.076<br>(0.083) | -1.499<br>(1.722) | 0.084<br>(0.083) | -1.911<br>(1.155) | -1.055*<br>(0.379) |
| Age | -0.003<br>(0.008) | 0.004<br>(0.004) | 0.031<br>(0.102) | 0.003<br>(0.004) | 0.057*<br>(0.029) | 0.138<br>(0.010) |
| $Age^2$ | -0.00004<br>(0.00007) | -0.00003<br>(0.00004) | -0.0007<br>(0.0015) | -0.00002<br>(0.00004) | -0.001*<br>(0.0002) | -0.0001<br>(0.0001) |
| Excon | 0.016*<br>(0.008) | 0.003<br>(0.004) | 0.066*<br>(0.033) | 0.002<br>(0.005) | 0.052<br>(0.039) | 0.021<br>(0.017) |
| Employd | 0.080<br>(0.049) | -0.004<br>(0.021) | 0.356*<br>(0.203) | -0.006<br>(0.021) | 0.313<br>(0.251) | -0.031<br>(0.056) |
| Drugs | 0.080<br>(0.050) | 0.013<br>(0.024) | 0.291<br>(0.197) | 0.013<br>(0.021) | 0.243<br>(0.228) | 0.089<br>(0.082) |
| Famcount | 0.019<br>(0.015) | -0.003<br>(0.006) | 0.072<br>(0.063) | -0.002<br>(0.006) | 0.058<br>(0.069) | -0.077*<br>(0.023) |
| Bond | -0.203*<br>(0.124) | ... | -1.015*<br>(0.556) | ... | -1.661*<br>(0.589) | ... |
| NOB | 287 | 1,344 | 287 | 1,344 | 2,311 | 2,311 |
| Cross Equation Correlation Coefficients | | | | | | |
| $r_{12}$ | | | | | 0.077<br>(0.528) | |
| $r_{13}$ , $r^*_{13}$ | | | | | -0.259<br>(0.538) | 0.274<br>(0.337) |
| $r_{23}$ | | | | | 0.385<br>(0.772) | |
| Predicted Average Failure To Appear Rate For Holdout Sample | | | | | | |
| | 0.171 | 0.155 | 0.146 | 0.152 | 0.093 | 0.231 |
| Predicted Number of Cases of Failure To Appear For Holdout Sample | | | | | | |
| | 789 | 290 | 501 | 587 | 148 | 1,770 |
| (Pr>0.2 Out Of 2027 Cases) | | | | | | |

to raise probability of failure to appear. Of course, the analysis presented above suggests that such results may arise due to upward selectivity bias on the estimated constant term in path A and downward bias in estimates of the path B constant. It is interesting that the single-equation probit results reverse the anticipated pattern. The estimated constant term of the path A probit equation is much smaller than that for path B. However, there are also very large differences in the estimated coefficients for age and age squared which may offset the differences in these constant terms.

The path A estimation results contain the variable Bond, the dollar amount of bail set and bond posted. This variable cannot be observed for those released on recognizance, and hence it does not enter the path B estimates. The estimated coefficient for Bond is consistently negative and significant. In the probit path A estimates, the estimated coefficient of Bond is numerically large and may have a substantial negative influence on the predicted probabilities of failure to appear obtained by using the estimated equation.

The single equation estimation results may be compared and evaluated by using the estimated coefficients to compute predicted probabilities of failure to appear using the holdout sample, as was done in the previous chapter for pretrial arrest. The bottom of Table VIII-2 shows results of these holdout sample predictions in the form of predicted average rates of failure to appear and predicted numbers of failure to appear out of the holdout sample of 2027 cases.

There is one special problem with estimation of failure to appear using the path A results that deserves special attention. Bond is not observed for persons released on recognizance, and some persons who had bail set were not able to post bond. This latter group were still treated as if they were able to secure release and no adjustment was made in their observed bail amount because the form of the path A failure to appear equation requires that bond be observed. Essentially, this is a conditional failure to appear equation, conditional on the accused obtaining release on bond.

Because the estimated coefficient of Bond is numerically large and statistically significant, the choice of a dollar bond for cases in which it is not observed is quite important. Clearly, failure to appear estimates can be made very small if large values of Bond are used. In the estimates of average probability reported in Table VIII-2, the value of Bond for cases where no bond was set is an estimated value derived from a statistical bond amount prediction equation. This equation was estimated by regressing bond amount on personal characteristics of the accused, including criminal history, using only cases where a bond amount was observed. Such estimates are themselves subject to selection bias but, given the limited use made of estimated bond amounts, no elaborate econometric adjustments were made to the estimation results.

The predicted average probabilities of failure to appear are quite counter-intuitive until one recognizes the anticipated influence of selectivity bias. For OLS estimates, the path A coefficients predict about 10% higher probability of failure to appear, 17.1% average probability vs. 15.5%. Taken literally, this could be interpreted as implying that setting bond raises the probability of failure to appear substantially. Differences in average probability of failure to appear based on the path A vs. path B probit results are very small, 14.6% for those released on bond vs. 15.2% for release on recognizance, but still suggest a slightly higher failure to appear rate for those released on recognizance. However, such results are quite consistent with the hypothesis that release on bail has no effect on failure to appear.

Table VIII-2 also shows the estimated coefficients obtained by estimating failure to appear equations for path A using trivariate probit to allow for possible selectivity bias in the three-stage decision process and for path B using bivariate probit to allow for selectivity bias in the two-stage process for those released on recognizance. The estimated constant term for path A is significantly less than that for path B and, as in previous single-equation results, the estimated coefficient of Bond is negative and statistically significant. The OLS and binary probit results in Table VIII-2 are conditional estimates of the probability of failure to appear for individuals on each path conditional on the selection rule used to divide the sample of accused persons. In contrast, the trivariate and bivariate probit results are unconditional results in which the estimated coefficients are adjusted for potential bias due to the selection rule which sends the higher-risk accused to the bond system.

These estimated equations using bivariate or trivariate probit techniques may be used to make unconditional forecasts of the expected rate of failure to appear that would occur if all accused were either released on recognizance or on bond. This was done using the holdout sample according to the special procedures, particularly those dealing with cases where no bond was originally set, described above.

The resulting average probabilities of failure to appear predicted for the holdout sample are shown at the bottom of Table VIII-2. The predicted average probability for path A with bail set is 9.3% compared to 23.1% for path B release on recognizance. These results contrast sharply with those for single-equation techniques where OLS results gave predicted probabilities higher for path A, and simple probit showed virtually no difference. Such differences were anticipated based on the likely influences of selectivity bias on single-equation estimates. The bivariate and trivariate estimation results, taken together, suggest that substantially higher failure to appear rates would be observed if all arrested persons were released on recognizance than if all were released on bond. The observed rate of failure to appear for those actually released, shown in Table VIII-1, was about 16.5% which reflects a mix in which the majority of releases were on recognizance. The estimated rates of failure to appear in Table VIII-2 are for the case in which all of the holdout sample is given a particular form of release, including individuals who were held previously.

An alternative measure of the effect of different release conditions as measured by different estimates of the failure to appear equation is found by examining the predicted number of cases of failure to appear at the bottom of Table VIII-2. These estimates are all obtained using the same holdout sample of 2,027 cases. As with the average probabilities, there are dramatic differences in the relative number of predicted failures to appear between paths A and B depending on the choice of estimation technique. Because the average probability of failure to appear is low, a predicted failure was associated with any case where the estimated probability was greater than 0.2 (compared to the 0.5 standard used for pretrial arrest estimates above). OLS estimates show far more predicted cases of failure to appear if everyone were released on bond than if they were released on recognizance. Probit estimates show a small reduction in failure to appear, 501 compared to 587 cases, if all arrested persons were released on bond. However the trivariate and bivariate estimates indicate very large reductions in predicted failure to appear if release on bond were universal compared to release on recognizance.

The absolute or numerical value of these estimates of cases of failure to appear is, of course, an artifact of the use of the 0.2 probability standard. But the estimates do indicate how different the implications of the different estimators are and how important it is to consider the potential effects of sample selection.

The trivariate estimates of path A and bivariate estimates of path B strongly suggest that release on bail does promote appearance compared to release on recognizance. This contrasts to the conclusions drawn using estimates from single equation models, particularly OLS, and with the simple observation of average rates of failure to appear for accused persons posting bond vs. those released on recognizance. The evidence from these estimates suggests that selectivity bias may be very strong in the pretrial release process, precisely because the judges and other actors use many characteristics that are difficult to observe in making release decisions and because they do succeed in differentiating between high- and low-risk cases. Thus it appears that the current system does select the lower-risk accused for release on recognizance.

The significant effect of bond on failure to appear found by comparing trivariate estimates of path A with bivariate estimates of path B could have been anticipated by inspecting the single-equation estimates of path A. Note that the estimated coefficient of Bond is consistently negative and statistically significant. This implies that raising Bond lowers the expected probability of failure to appear among those accused actually securing release. Clearly, such a result suggests that bond has an effect on appearance conditions and this incentive effect should be to reduce the probability of failure to appear just as shown in the empirical results. Thus, the estimated coefficients obtained using single-equation techniques that only produce conditional estimates imply that bond setting lowers failure to appear.

Yet comparison of OLS and probit estimates of path A vs. path B fails to reveal a significant deterrent effect associated with release on bail vs. recognizance. One interpretation of this is that the influence of selectivity bias on estimates of the constant term and estimated coefficients of other variables, particularly age and age squared, offsets the deterrent effect suggested by the negative and significant coefficient of Bond.

As noted at the beginning of this section, the relationship between release on bond and the subsequent rate of failure to appear has been the object of some controversy in recent years. Reforms in the pretrial release system have suggested that use of bail be reduced and, along with it, the role of the bondsman. Unfortunately, the data do not permit us to identify precisely the cases in which bond is posted with the aid of a bondsman or the terms under which the bond contract is written. Interesting questions concerning the relative effectiveness of cash vs. deposit vs. surety bond could not be analyzed because sample sizes for these respective types of bond were too small.

One interesting result is the contrast between the importance of sample selection based on the above analysis and the lack of significance of estimates of the cross-equation correlation terms. The signs of the estimated correlations in Table VIII-2 are also not in agreement with expectations. While $r_{13}$ and $r^*_{13}$ are approximately equal numerically and opposite in sign as anticipated, it was expected that $r_{13}$ would be positive and $r^*_{13}$ negative. This continues a trend in which estimates of cross-equation correlation coefficients have low levels of statistical significance and sometimes do not have the anticipated signs.

## IX. EVALUATION OF "REDUCED FORM" PROBIT ESTIMATION TECHNIQUES

Thus far it has been argued that correction of estimates of behavioral equations which predict pretrial arrest or failure to appear for selectivity bias should be accomplished through use of a multivariate probit estimation technique with as many variables as there are stages to the selection process. Thus, bivariate probit was used for two-stage processes, and trivariate probit was used for three-stage systems. Given that some criminal justice systems involve four, five, or more stages, this implies that increasingly elaborate estimation routines be used. An alternative view is that any multi-staged selection process may be collapsed into two stages forming what will be termed, following Rhodes [1984], a "reduced form" probit model which may be estimated using bivariate probit estimation techniques.

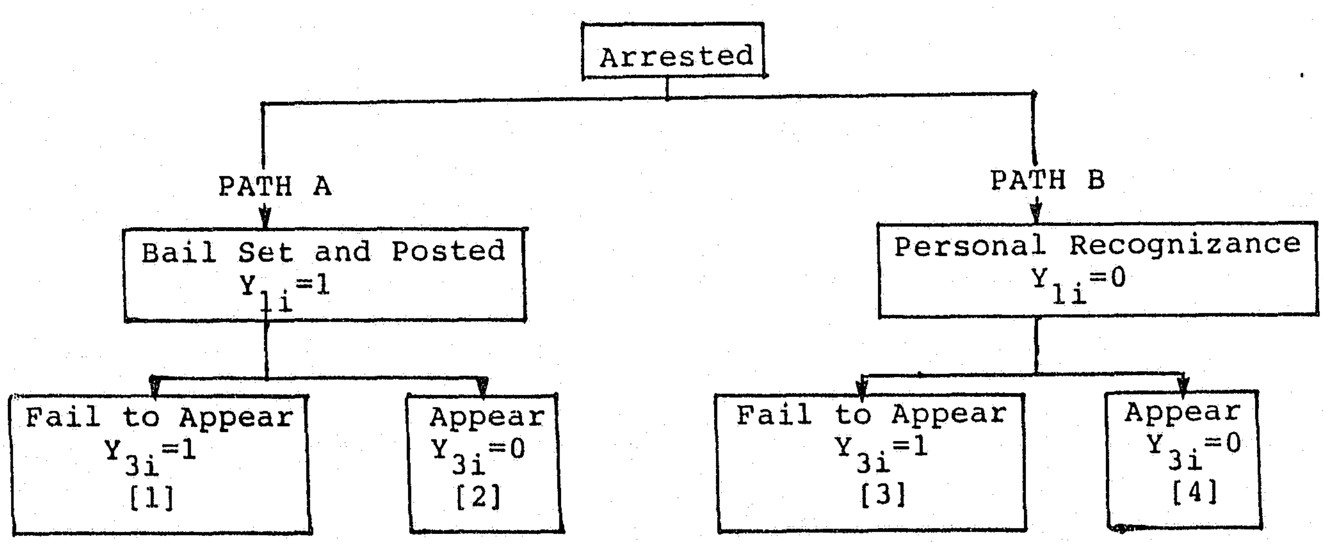
This section examines the argument for a "reduced form" probit approach by applying the technique to the three-stage process forming path A of the failure to appear model of the previous section. The reduced form probit estimates, constructed using bivariate probit techniques, can then be compared to the trivariate probit and simple probit estimates obtained above.

This test does not constitute a proof of the unbiasedness of the reduced form approach because it can be shown that this is not equivalent to the trivariate maximum likelihood estimator. Rather, it can indicate, for a particular estimation problem, the amount of increase in selectivity bias problems which one encounters when trading off the simplicity of a reduced form model for a full structural multivariate model of the system. One can conjecture that the problems with the reduced form approach would only increase as the number of stages being collapsed into a single stage increased. In this example, two stages are collapsed as shown in Figure IX-1.

Comparison of Figure IX-1 with the full system in Figure VIII-1 indicates that one group, those not released, has been lost to the analysis in the process of collapsing from three to two stages. The first stage decision now sets $Y_{1i}=1$ when bail is set and posted or when, under the three-stage notation, $Y_{1i}=1$ and $Y_{2i}=1$. There is no change in the condition for $Y_{1i}=0$. Hence, compared with the three-stage process, the observations for which $Y_{1i}=1$ and $Y_{2i}=0$ have been eliminated from the analysis.

FIGURE IX-1

REDUCED FORM REPRESENTATION OF FAILURE TO APPEAR IN FIGURE VIII-1



The reduced form model can be represented in terms of equations in the following way.

$$\text{(IX-1)}\quad Y_{1i} = G_1 + Z_{1i}g_1 + e_1$$
$$Y_{3i} = G_3 + Z_{3i}g^3 + e_3$$

As noted above, the outcome $Y^*_{1i}=1$ is observed for cases in which release is achieved after bail is set and bond posted and $Y^*_{1i}=0$ for release on recognizance. Cases in which bail was set and not posted are dropped from the analysis. Thus, the problem of unobservable outcomes tends to force truncation of the sample when the system is collapsed to a reduced form. $Y^*_{3i}=1$ is observed when the accused fails to appear and $Y^*_{3i}=0$ when appearance is made. The equations IX-1 may be estimated using maximum likelihood bivariate probit techniques on the sample of released persons. The estimation technique will correct coefficient estimates for correlation between $e_1$ and $e_3$, $r^*13'$ which should again be positive by the arguments made above.

Estimation results for the failure to appear equation from OLS, binary probit, reduced form bivariate probit, and trivariate probit all for path A are presented in Table IX-1. Examination of the estimated coefficients indicates that the reduced form bivariate probit results are quite close to those from the trivariate probit. Indeed, it appears that, from the point of view of classification of individuals, the two estimated equations would produce quite similar results. However, simple comparison of individual estimated coefficients can be misleading, given the non-linear nature of the relationship between the value of the probit function and the estimated probability of failure to appear.

TABLE IX-1
OLS, PROBIT, "REDUCED FORM" PROBIT AND TRIVARIATE PROBIT ESTIMATES OF DETERMINANTS OF FAILURE TO APPEAR
Third Equation Of The Model: Failure To Appear Path A
Estimated Coefficients With Standard Errors In Parentheses
* Indicates Estimated Coefficient Significant At 10% Level

| Independent Variables | OLS | Probit | Reduced Form Bivariate Probit | Trivariate Probit |
|---|---|---|---|---|
| Constant | 0.200<br>(0.204) | -1.499<br>(1.722) | -2.085<br>(1.886) | -1.911<br>(1.155) |
| Age | -0.003<br>(0.008) | 0.031<br>(0.102) | 0.059<br>(0.102) | 0.057*<br>(0.029) |
| $Age^2$ | -0.00004<br>(0.00007) | -0.0007<br>(0.0015) | -0.0011<br>(0.0014) | -0.001*<br>(0.0002) |
| Excon | 0.016*<br>(0.008) | 0.066*<br>(0.033) | 0.061*<br>(0.037) | 0.052<br>(0.039) |
| Employd | 0.080<br>(0.049) | 0.356*<br>(0.203) | -0.326<br>(0.229) | 0.313<br>(0.251) |
| Drugs | 0.080<br>(0.050) | 0.291<br>(0.197) | 0.312<br>(0.204) | 0.243<br>(0.228) |
| Famcount | 0.019<br>(0.015) | 0.072<br>(0.063) | -0.066<br>(0.076) | 0.058<br>(0.069) |
| Bond | -0.203*<br>(0.124) | -1.015*<br>(0.589) | -1.003*<br>(0.644) | -1.661* |
| NOB | 287 | 287 | 2,311 | 2,311 |
| Cross Equation Correlation Coefficients | | | | |
| $r_{12}$ | | | | 0.077<br>(0.528) |
| $r^*_{13}$, $r_{13}$ | | | 0.186<br>(0.603) | -0.259<br>(0.538) |
| $r_{23}$ | | | | 0.385<br>(0.772) |
| Predicted Average Failure To Appear Rate For Holdout Sample | 0.171 | 0.146 | 0.117 | 0.093 |
| Predicted Number of Failures To Appear For Holdout Sample (Pr>0.2 Out Of 2027 Cases) | 789 | 501 | 265 | 148 |

The bottom of Table IX-1 contains estimates of the average probability of failure to appear and of the predicted number of cases from the holdout sample with an estimated probability of failure to appear greater than 0.2. Compared to either the OLS or binary probit estimates, the reduced form bivariate probit results are closest to the trivariate estimates. The average percentage of failure to appear predicted for the holdout sample using reduced form probit is 11.7% which is far closer to the trivariate results of 9.3% than either the OLS estimate of 17.1% or the simple probit at 14.6%. The predicted number of cases for the reduced form probit, 265, is also fairly close to the 148 predicted using coefficients estimated by the trivariate probit. The 501 and 789 case estimates obtained from simple probit and OLS are quite high.

Overall, in terms of relative error magnitudes, the reduced form probit technique appears to be a clear improvement on single-equation methods. Given the current state of research and the widespread use of single-equation approaches, one could argue that widespread use of bivariate probit estimators for reduced form models of behavior in the criminal justice system would be a big improvement. However, only one case has been examined here and the reduced form results are, as one might have expected based on simple intuition, intermediate between single-equation approaches and the full trivariate results. Also, the case examined here is most favorable to the reduced form approach because only one stage has been collapsed: trivariate has been compressed into a bivariate probit model.

Certainly, greater numbers of stages might be collapsed but perhaps at greater cost in terms of bias.

Two additional interesting results of this particular test deserve attention. First, the estimate of $r^*_{13}$ from the reduced form bivariate probit has the expected positive sign in contrast to the negative sign on the $r_{13}$ obtained from the trivariate probit estimator. However, neither estimate is statistically significant and the high standard errors for the cross equation correlation coefficients remain a disappointment. Although compared to the observations of Schmidt [1984], who reports that bivariate probit results in the literature have problems with $r_{12}$ falling below -1.0 or rising above 1.0, the standard error problems experienced here are small.

Both the reduced form bivariate probit and the trivariate probit estimates of failure to appear suggest that forcing all arrested persons through a system of release on bail would result in lower rates of failure to appear. Thus, both estimators produce similar general implications for policy toward reducing failure to appear. Of course, the results are not identical and the similarity may be an artifact of this particular application because there is no reason, in theory, for the reduced form bivariate probit formulation to produce unbiased estimation results.

# X. SUMMARY AND CONCLUSIONS

The preceding chapters have developed a theoretical analysis which suggests that conventional statistical models of behavior in the pretrial justice system may produce biased estimates. This theoretical point follows if actors in the justice system, particularly judges and magistrates, select accused persons for differential treatment based on characteristics which are not directly observable. It is possible to use multivariate probit techniques to eliminate the selectivity bias due to the differential treatment of accused persons. Essentially, the analysis makes clear a fundamental problem in trying to develop better classification systems for accused persons or trying to evaluate the efficacy of current treatment strategies. The current methods of classification, particularly if they are effective, produce selected samples. The position of an accused person in the pretrial justice system is based on a prior assessment of the risk of misconduct. Such selection produces very heterogeneous groupings of persons in different treatment groups and makes econometric estimation of the behavior of these groups most difficult.

Based on suggestions by Lee [1984], a multivariate probit estimation technique was implemented to allow estimation of relationships in selected samples drawn from the pretrial justice system. A test for selectivity bias was conducted by estimating a variety of models using conventional single-equation techniques and comparing them to the multivariate probit results. In general, the differences in statistical results are in the direction and of the type which would be expected if there were substantial selection bias introduced by behavior of various actors in the pretrial justice system.

The quantitative results presented here are based on data from the pretrial justice system operating in Washington, D.C. and on the particular mix of arrested persons found in this area. It is possible that differences among jurisdictions are large enough so that these results would not generalize across areas. For example, the degree of selectivity bias depends on the effectiveness of judges and magistrates to differentiate among accused persons and to detain those highest risk cases. If this assignment process were random, then no selectivity problems would arise. Clearly, the results indicate that classification of higher risk individuals into restricted release groups is quite common and this promotes selectivity problems. Of course, from a justice system operation viewpoint, such successful classification is laudable. But, as noted above, the problem is that good classification by judges and magistrates tends to produce selected data that creates problems for econometric analysis.

Other aspects of the criminal justice system in the District of Columbia might have a significant influence on the results. The use and sophistication of bondsmen may vary geographically. This influences the degree of selection occurring when individuals, for whom bail was set, either fail or succeed in securing their release. Some factors may appear to have a potentially significant influence on the results but not be very important for the type of tests conducted here. For example, the ability of pretrial arrest to serve as a proxy for pretrial crime will certainly vary geographically with factors such as arrest and clearance rates. Also, the emphasis placed on prevention and detection may vary by type of charge. However, it is not clear how such differences in procedures and results would cause the tests for selectivity to vary across locations.

Despite these limitations, the results of the analysis have important implications for criminal justice decision makers who are involved in the classification of defendants on the basis of release risk. The results obtained using both bivariate and trivariate probit estimators to correct for selectivity bias differ significantly and systematically from classification results obtained using conventional single-equation approaches that are subject to bias. It is important to note that the differences were systematic in that the direction of bias was anticipated before the estimates were made. This provides particularly strong evidence that the differences in estimation results are due to selectivity bias.

The general pattern of selectivity bias in the empirical results is easily characterized. With estimates of determinants of both pretrial arrest and failure to appear, the conventional single-equation approaches tend to produce estimates of misconduct which are too low when the data used for the estimation are based on persons given unconditional release. This result was expected because persons given such release are expected to be better risks and have lower probability of misconduct. Conversely, those given more restrictive release, particularly those released on bond, have higher expected probability of misconduct and this, presumably, accounts for the restrictions on their release. The conventional single-equation estimates of probability of misconduct when the observations are drawn from these groups produce predicted probabilities which are too high.

Thus, the general pattern of observed and expected selectivity bias is that conventional single-equation models give conditional estimates based on the data used for the estimation. If the data are based on a group selected because they reflect good risks, then conditional estimates of misconduct will be below those that would follow from an unconditional estimate which used data on all accused persons. One might well ask: why not use data on all accused? There are two problems with this approach. First, not all accused persons are released and given the opportunity for pretrial misconduct. This is the problem of partial observability. Second, accused persons are released under different conditions, and their subsequent behavior is based both on their underlying riskiness and on the incentive effects added by the release conditions. Again this may be thought of as a problem of partial observability because not all persons are given release under identical circumstances - yet this is the experiment that would be needed to produce an unconditional data set and allow unbiased estimation using conventional approaches.

The nature of the bias in estimation results obtained using conventional approaches on selected data is demonstrated by estimating pretrial arrest or failure to appear equations. Such equations would ordinarily be used to classify accused persons or to determine the consequences of releasing such persons under various conditions. The estimated equation is used to construct a predicted probability of pretrial arrest or failure to appear that ranges from zero to unity. In order to demonstrate differences between conventional and multivariate probit approaches to such estimation, predictive tests were made using a holdout sample with characteristics similar to the initial data used for estimation. The expected probability of pretrial arrest or failure to appear could be computed for each individual in the holdout sample.

One comparison between classification equations based on conventional vs. multivariate probit estimates was based on the average predicted probability of pretrial arrest or failure to appear for individuals in the holdout sample. As expected, the conditional predictions obtained from conventional techniques were below the unconditional estimates from multivariate probit models when data sets consisted of persons selected as "good" risks. Conversely, for data sets consisting of the highest risk cases, the estimates obtained from conventional techniques produced predicted probabilities which were above those obtained using multivariate probit. If classification schemes were instituted based on estimates obtained using conventional approaches, the judge or magistrate making release decisions would face estimated probabilities of pretrial misconduct which tended to underestimate risk for persons released unconditionally and overestimate risk for those on whom the most significant conditions were placed.

Using the holdout sample, the number of cases with a predicted probability above a given standard (0.5 was used for pretrial arrest and 0.2 for failure to appear) was compared. Conditional estimates using single-equation approaches generally gave quite low estimates of expected pretrial arrest or failure to appear when the sample was for cases given relatively unrestricted release conditions compared to the unconditional estimates obtained from multivariate probit. Conversely, the numerical estimates of pretrial misconduct for the conditional estimates were significantly higher than unconditional estimates when data from those given strict release conditions were used. In some cases the differences in predicted numbers of cases were very large and the implications for efficacy of different release strategies substantially affected. For example, conditional estimates give the impression that release on bond has little or perhaps even negative effect on failure to appear. However, the unconditional estimates from multivariate probit suggest that release on bail does act as a deterrent to failure to appear, but not to pretrial arrest.

Taken together, the exercises in which conditional estimates from conventional approaches are compared to unconditional estimates from the multivariate probit estimator developed here suggest that selection bias is substantial in the conditional estimates of behavior in the pretrial justice system. Decisions on classification criteria, particularly on the overall level of expected risk of misconduct, should be made using unconditional estimates. The multivariate probit techniques developed in connection with this report can provide

Another comparison between conditional results from conventional estimation techniques and unconditional results from multivariate probit was based on the total number of predicted cases of pretrial arrest or failure to appear in the holdout sample. In order to make a prediction of number of cases, some probability standard must be adopted. For example, if 0.5 is the standard, then all cases in the holdout sample with an estimated probability of pretrial arrest equal to or greater than 0.5 would be predicted to experience arrest and cases with a predicted probability below 0.5 would be classified as non-arrest predictions. The probability standard of 0.5 is important because it makes explicit a standard of expected dangerousness or flight risk which is being used to justify release conditions.

As the probability standard falls toward 0, the number of cases of predicted pretrial misconduct increases. The standard, together with the classification equation used to estimate pretrial misconduct, makes clear and explicit the policy tradeoff between expected misconduct and the number of persons whose release is restricted. Given limitations of capacity to detain accused persons, it is important to be able to predict the number of persons who would be detained if a particular standard of expected probability of pretrial misconduct were adopted.

The results obtained from comparisons of number of predicted cases of pretrial arrest or failure to appear between the conventional and multivariate probit estimation techniques follow those discussed above for the average probabilities.

such unconditional estimates for classification and policy development purposes.

# APPENDICES

# APPENDIX A

## STATISTICAL ANALYSIS OF RELEASE DECISIONS

In the process of constructing estimates of the probability of pretrial arrest and failure to appear, estimates of behavorial equations characterizing the pretrial release process were developed. This section considers, specifically, estimates of the release decision which were made in conjunction with the first stages of the various models considered. Two types of first-stage release decisions were estimated in the context of the analysis. First, the decision to set sufficiently strict release conditions so that the accused was held was examined in the first stage of the estimation presented in Chapters VI and VII. See Figures VI-1 and VII-1 for a more revealing insight into the structure of these models. Second, the probability of setting a financial condition, i.e. setting bond, was examined. Chapter VIII presents a three-stage model in which the first-stage bail-setting decision considers the probability of setting a financial condition, bail, as opposed to release on recognizance.

While the multi-stage nature of the pretrial release process results in exposure of selected samples to different forms of treatment, the entire sample of arrested persons is exposed to an initial release decision of the type discussed here. Therefore, there is no problem of partial observability or of selectivity bias. The parameter estimates obtained using single-equation techniques, such as simple probit, should produce unbiased estimates of the probability-of-release equation. This proposition was tested by estimating release equations using single-equation techniques and comparing the results with estimates from bivariate or trivariate probit. As expected, there were no significant differences in the parameter estimates. This may be seen as a confirmation of the precision of the multivariate probit estimation routines which were compared to results from proben single-equation probit estimation packages such as CRAWTRAN.

Table A-1 contains the estimation results for the release and bail-setting equations discussed above. The release without nonfinancial conditions equation, otherwise known as the probability-of-release equation is presented first in the table. This is really a model of both judicial behavior and of the accused and bondsman. In order to secure release for those given a financial condition, either they must be willing and able to post bond or be able to convince a bondsman to post bond. In addition, the D.C. Pretrial Services Agency also influences the release decision. As might be expected, increasing prior experience with the criminal justice system tends to reduce the probability that the accused will secure release. For example, the estimated coefficients of Excon (number of prior convictions), Pendcase (number of cases pending at arrest), and Parole (a dummy variable equal to one if the accused is on parole) are all negative and statistically significant. These are the principal sources of objective information on the frequency with which the defendant has encountered the criminal justice system in the past and judges or magistrates apparently take these factors seriously. The size of the estimated

TABLE A-1

PROBIT ESTIMATION RESULTS FOR RELEASE EQUATIONS

Estimated Coefficients With Standard Errors In Parentheses
* Indicates Significance At The 10% Level

| Independent Variable | Release Equation | Bail Set Equation |
|---|---|---|
| Constant | 0.604* | 0.187 |
| | (0.179) | (0.400) |
| Excon | -0.142* | 0.021 |
| | (0.013) | (0.016) |
| Pendcase | -0.302* | 0.002 |
| | (0.048) | (0.003) |
| Parole | -0.582* | 0.036 |
| | (0.076) | (0.033) |
| Probation | 0.392* | 0.057* |
| | (0.082) | (0.021) |
| Confidence | -0.269* | -0.181 |
| | (0.130) | (0.146) |
| Violent | -0.298* | |
| | (0.132) | |
| Drugs | 0.441* | -0.106 |
| | (0.110) | (0.101) |
| Larceny | 0.254* | -0.654* |
| | (0.117) | (0.157) |
| Prostitution | 0.786* | |
| | (0.239) | |
| Weapons | 0.441* | |
| | (0.186) | |
| Posesscrim | 0.432 | |
| | (0.199) | |
| Miscrim | 0.105 | |
| | (0.113) | |
| Age | | -0.028* |
| | | (0.015) |
| $Age^2$ | | 0.00022 |
| | | (0.00015) |
| Employed | | -0.072 |
| | | (0.093) |
| Homeowner | | -0.170 |
| | | (0.103) |

coefficients for Excon, Pendcase, and Parole also indicates that these are very important influences on release probability. All other things being equal, a defendant with 2 prior convictions with a pending case currently on parole has a probability of release that is about 33 percentage points lower than for a defendant with no prior criminal history.

The type of charge at arrest also influences the probability of release. Curiously, both confidence and violent charges, where violent includes murder, rape, kidnapping, and robbery, are associated with lower rates of release. Conversely, prostitution has a very strong positive association with release probability. Having a most serious charge of larceny, weapons, possession of criminal implements, or drugs violation is associated with a significant increase in the probability of pretrial release. Demographic and economic characteristics of the defendant are omitted from this equation because their estimated coefficients were found to have very low levels of significance in previous econometric testing.

The second type of release decision studied was the probability of setting bail, i.e. of setting a financial condition for release. The second column of Table A-1 shows the estimated coefficients from a bail-setting equation which was part of the first stage of the failure to appear model discussed in Chapters VIII and IX. Given that the setting of strict financial conditions is viewed as a means for lowering release probabilities, it was expected that defendants with the worst histories of criminal behavior would be most likely to have bail set. Descriptive data presented in Chapter VIII confirmed this expectation.

The estimated coefficients for the probability-of-bail-setting equation in Table A-1 show uniformly positive effects of the criminal history variables, Excon, Pendcase, Parole, and Probation. This confirms the expectation that prior criminal history is an important influence on the decision to set a financial condition. Note that the standard errors of the estimated coefficients are rather large and that only the estimated coefficient of Probation is significant at the 10% level. Generally, the standard errors in this bail-setting equation are large. This reflects the difficulty encountered in accounting for judicial bail-setting behavior in terms of readily observed characteristics of the accused. Presumably, this bail-setting behavior is more systematic and perhaps it is based on some unobserved factors which are correlated with the characteristics in the equation, thus accounting for the large standard errors. This is precisely the type of situation in which the potential for selectivity bias in the latter stages of estimation of the sequential decision system processing accused persons is large. Of course, Chapter VIII found that such selectivity bias did exist in failure to appear equations estimated using conventional techniques.

Many type of crime variables were tried in the bail-setting equation but they were generally nonsignificant, with the exception of larceny which has a large negative estimated coefficient. Economic variables such as Employed and Homeowner, which presumably reflect higher levels of income and/or wealth, might be thought to influence the setting of a financial condition. Both had negative estimated coefficients, with Homeowner just below the standard for significance at the 10% level. This is very weak evidence of the use of financial conditions on defendants who are less likely to be able to meet them. Demographic variables were not significant in this equation except for age which had a surprising negative and significant coefficient. Given that age has a nonlinear effect, because there is an age squared term, this result may simply reflect a lower probability of bail-setting for significantly older defendants.

The probability-of-bail-setting equation is potentially very important because financial conditions are an important means for lowering the probability that the defendant secures release. The low predictive power of the available information on the accused in this equation indicates that more detailed attention to the determinants of bail-setting is in order.

## APPENDIX B

## PROBABILITY OF SECURING RELEASE AFTER BAIL IS SET

In the models of failure to appear examined in Chapters VIII and IX, the second-stage decision in the pretrial release process involved the posting of bond by the accused conditional on bail having been set. Bond-posting is only observed for those defendants who have bail set. Thus, for most defendants, the decision to post bond is never observed. In this section, the probability of posting bail is analyzed and conditional estimates from single-equation probit are compared to the estimates from the trivariate probit equation used to study the three-stage pretrial release system which involved the probability of setting bail at the first stage, posting bond at the second stage, and failing to appear at the final stage. The results obtained for the probability-of-posting-bond equation are rather surprising. This issue has not received extensive formal econometric analysis but the results reported here indicate that it may be worthy of further study.

The probability of posting bond should depend on the resources of the accused, the aversion to spending time in prison awaiting disposition, and the decision of the bondsman to cooperate with the accused by posting the bond. Recall that, in the failure to appear equation, the amount of bond posted acted as a powerful deterrent which lowered the probability of failure to appear. Based on these considerations, a variety of variables were tested as possible arguments of a probability-of-posting-bond equation. Generally, the results were disappointing and it was difficult to find variables which had the expected sign and significance in a probability-of-posting-bond equation.

Results of single-equation probit estimates, made using only observations on defendants for whom bail was set, are shown in the first column of Table B-1. These estimated coefficients are conditional on the accused having bail set and presumably refer to a selected sample of defendants whose criminal careers have been rather extensive. The estimated coefficients were largely non-significant even after variables with t-ratios below 1.0 were eliminated.

It is rather surprising to note that defendants with more extensive criminal histories, as indicated by the magnitude of variables Excon (number of prior convictions), Parole (dummy variable for accused persons on parole), and Probation (dummy variable for defendants on probation) all have positive estimated coefficients and the latter two variables are significant. The estimated coefficients of these criminal history variables are not large. Indeed the partial effect of being on probation or parole on the probability of posting bond successfully is only about two percentage points. It is most surprising that Bond, the dollar amount of bail set, has a positive and significant effect on the probability of posting bond. This may reflect larger bonds being set for persons better able to post bond. The estimated coefficients of Employd (a dummy variable equal to one if the accused is employed) and Ownrent (a dummy variable equal to one if the accused is an owner or renter) are both negative. Again, one would imagine that persons holding jobs or homeowners would be better able to meet bail requirements.

TABLE B-1
SINGLE EQUATION PROBIT AND TRIVARIATE PROBIT ESTIMATES
OF THE PROBABILITY OF POSTING BOND

Estimated Coefficients With Standard Errors In Parentheses
* Indicates Significance At The 10% Level

| Independent Variables | Probit | Trivariate Probit |
|---|---|---|
| Constant | 0.210<br>(0.045) | -0.619<br>(0.574) |
| Excon | 0.007<br>(0.005) | 0.032<br>(0.023) |
| Employed | -0.037<br>(0.030) | -0.093<br>(0.094) |
| Ownrent | -0.045<br>(0.033) | -0.175*<br>(0.106) |
| Parole | 0.032*<br>(0.015) | 0.129*<br>(0.061) |
| Probation | 0.039*<br>(0.019) | 0.089*<br>(0.035) |
| Bond | 0.149*<br>(0.065) | 0.240<br>(0.212) |

The second column of Table B-1 shows the unconditional coefficient estimates obtained from the trivariate probit estimator. These coefficients reflect the probability that any defendant would succeed in posting bond if all accused had bail set. Surprisingly, the estimated coefficients from this unconditional model are similar to those for the conditional probit model. The three criminal history variables, Excon, Parole, and Probation, all have positive coefficients and the latter two are statistically significant. The strange single-equation probit result in which increasing bond amount raised the probability of posting bond is replaced by a trivariate probit estimate that is positive but nonsignificant. Again, the positive coefficient for Bond is surprising, even if it is nonsignificant. Finally, Employd and Ownrent, the two variables reflecting income, have negative estimated coefficients, significant in the case of Ownrent. This is most unusual, particularly the Employd variable. Based on theory, one would expect employed defendants to have the most resources available to post bond and to have the largest losses from being detained prior to disposition.

Overall, the results presented here contrast sufficiently with expectations about the incentives and ability to post bond to warrant further study. It may be that bond amounts and terms are adjusted so that those with greater ability to pay face larger bail amounts. But it is not clear why defendants with jobs and who are owner-occupants should not have an advantage in posting bond. Surely, additional research on such questions should be encouraged.

## APPENDIX C

## DOCUMENTATION FOR PROCESSED DATA ON PRETRIAL MISCONDUCT

The initial data source on pretrial misconduct was a data tape obtained from the Washington, D.C., Pretrial Services Agency. This tape which had a standard label, "BAILDANN," contained information on the entire population of arrest incidents for Washington, D.C., during the January 1980 to December 1982 period. The basic information contained in the arrest records on this tape is identical to that found in the first 344 card columns of the tape documentation supplied below.

For purposes of the Classification Systems for the Accused Project, this basic data on arrest incidents was sampled and processed in a number of important ways to create the data set documented below, which is available on request accompanied by a blank, initialized standard label tape 9 track suitable for IBM equipment. Each of the steps in the processing and sampling is described in turn below, followed by a complete data record documentation.

First, cases in which the arrest did not result in booking for a local crime were dropped by eliminating all cases in which the "RELEASE" variable was coded 14, no paper, or 30, turned over to. The no paper category means literally that no papers were filed and the accused was released without being booked or charged. It was felt that such cases should not count as arrest incidents, and particularly that they would be a misleading indication of pretrial crime. If the accused was turned over to another jurisdiction, this is an indication of an arrest for a previous incident and it is not likely that there would be an opportunity for local pretrial crime in such cases. This category was small, less than 30 cases, while there were several hundred no paper cases.

Next, a period of time which was called the "arrest window" was selected. The period January 1, 1981 to July 1, 1981 was selected in order to allow sufficient time to observe both pre- and post-arrest behavior in the data. The data were then sorted by police identification number, PDID, which is unique for each person arrested. Then, each arrest which occurred in the arrest window was allowed to create an "arrest record" in which information on the current arrest was combined with information on: prearrest arrests, all arrests occurring before the current arrest; pretrial arrest, all arrests occurring during the pretrial or predisposition period for the current arrest charge; and post-trial arrest, all arrests occurring after disposition of the current arrest charge. Thus, each arrest in the arrest window divided the January 1980 to December 1982 period into three segments: prearrest, pretrial, and post disposition. Information on prearrest arrests, pretrial arrests, and post disposition arrests was added to each arrest record to create a single record for each arrest in the arrest window which gave a 3-year criminal justice system history for the individual.

The above procedures created a data set consisting of the entire population of arrest cases which resulted in charges being filed for January through June 1981. Note that an individual could be arrested several times during this period. These subsequent arrests would be counted as pretrial or post disposition arrests on the initial arrest record for the period and they would be counted as prearrest arrests on subsequent records. Thus, "active" arrestees would account for an elevated proportion of the sample. In effect, the frequency of rearrest would influence the sample contents and statistical inference about pretrial arrest in such a sample would be difficult.

This problem with the arrest-based sampling was eliminated by extracting a sample consisting of the first arrest in the arrest window for each arrested person. This "person-based" sample is the population of persons arrested during the arrest window period and the frequency of rearrest does not influence the number of arrest records in the sample. Results based on this "person-based" sampling should only be used for making inferences about the population of arrested persons, not about a population of arrest incidents. This is a subtle but important issue that has been ignored in statistical analysis of similar data sets.

The person-based sample yielded 4,253 cases. These were divided randomly, using the last digit of the police identification number, into a 60% sample of 2,311 cases which were used for econometric analysis and a 40% holdout sample of 1,942 cases used to generate implications of alternative model estimates.

The table below indicates the basic format of the data record along with variable names and descriptions of the way in which the variables are coded.

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 1-4 | PSANO1 | f4.0 | Pretrial Services Agency ID Code. |
| 5-8 | PSANO2 | f4.0 | |
| 9 | CASEST | f1.0 | Case status 1=open, 2=closed, 3=appeal |
| | | | Arrest Date |
| 10-11 | FIYR | f2.0 | Year of Papering with Court |
| 12-13 | FIMO | f2.0 | Month of Papering with Court |
| 14-15 | FIDY | f2.0 | Day of Papering with Court |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 16-17 | FDISP | f2.0 | Final disposition<br>0=Case Open, 1=No Paper, 2=Nolle,<br>3=Dism W Prejudice, 4=Ignored by GJ,<br>5=GJ Abatement, 6=Dropped No Prosec,<br>7=MJOA, 8=Not Guilty, 9=NG Reas Insanity,<br>10=Not Comp to Stand Trial, 11=No Contest,<br>12=Deceased, 13=Other No Sent,<br>14=Dism WO Prejud, 21=Security Forfeited,<br>22=Fined, 23=Fine or Days,<br>24=Sent to Time Served,<br>25=Time Less Than 1 Day, 26=1 DY to 1 Year,<br>27=Over 1 YR to 5 Years, 28=Over 5 to 10 YRS,<br>29=Over 10 Years, 31=ESS No Prob,<br>32=RVTDS Removal, 33=Prob-Unsup,<br>35=Prob up to 1 Year, 37=Prob 1-5 Yrs,<br>39=Prob over 5 Yrs, 41=Work Release,<br>42=Work Rel-Prob, 43=FYCA-Prob, 44=FYCA-B,<br>45=FYCA-C, 46=FYCA-D, 47=NARA,<br>48=Other Sentence, 49=Extradited,<br>50=No Probable Cause, 51=Prob WO Judgment |
| 18-23 | TIMET | F6.0,6.2 | Time to Disposition of the Case<br>in Days |
| 24-27 | CASETIME | f4.0 | " " " " " " " " |
| 28-29<br>30-31<br>32-33 | FLTREC1<br>FLTREC2<br>FLTREC3 | f2.0<br>f2.0<br>f2.0 | Flight Recommendation Reasons-New Scheme<br>Table of Outcomes for FLTREC1-FLTREC3<br>1=Straight PR, 2=PR-Appearance,<br>3=No Safety, 4=No Appearance,<br>5=Warrant-Detainer, 6=Mo Hospital,<br>7=Name Identity, 8=Hold WO Bond,<br>9=Address Problem, 10=No Interview,<br>11=No Paper, 12=Nolle, 13=Dismissed,<br>14=RVTDS, 15=Unable, 16=TOT, 17=Contempt,<br>18=Solve For Under Sent, 19=Missing, 20=??? |
| 34-35<br>36-37<br>38-39 | SAFREC1<br>SAFREC2<br>SAFREC3 | f2.0<br>f2.0<br>f2.0 | Safety Recommendation-New Scheme<br>Table of Outcomes for SAFREC1-SAFREC3<br>1=Straight PR, 2=No Safety, 3=PR-Safety,<br>4=E Hearing Prob, 5=E & A Hearing Prob,<br>6=E Hearing Parole, 7=E & A Hearing Parole,<br>8=E Hring Prob & Par, 9=E & A Prob & Par,<br>10=A Hearing-Dang, 11=A Hearing-Witness,<br>13=Warrant-Detainer, 14=MO Hospital,<br>18=No Interview, 19=Contempt, 21=Missing,<br>22=??????? |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 40-41 | RFLTCD1 | f2.0 | Flight Condition Recommended-New Scheme |
| 42-43 | RFLTCD2 | f2.0 | Table of Outcomes for RFLTCD1-RFLTCD5 |
| 44-45 | RFLTCD3 | f2.0 | 1=Interstate Superv, 2=Send Notice, |
| 46-47 | RFLTCD4 | f2.0 | 3=Live At, 4=Third Party-Person, |
| 48-49 | RFLTCD5 | f2.0 | 5=Provide PSA Address, 6=Rpt to PSA in Person, 7=Rpt to Prob upon Rel, 8=Rpt to Par upon Rel, 9=Custody, 10=Halfway House, 11=Work Rel From Jail, 12=Surrender Passport, 13=Person for Notice, 14=Live At, 15=Rpt to Military Org, 16=Provide PSA Address, 17=Rpt to PSA by Phone, 18=Rpt to Sent Judge, 19=Stay in DC Area, 20=Surrender Passport, 21=Interstate Superv, 22=Maint Psych Treat, 23=Missing, 24=??????? |
| 50-51 | SCDRECD1 | f2.0 | Safety Condition Received-New Scheme |
| 52-53 | SCDRECD2 | f2.0 | Table of Outcomes for SCDRECD1-SCDRECD5 |
| 54-55 | SCDRECD3 | f2.0 | 1=Speedy Trial, 2=24HR Residen Custody, |
| 56-57 | SCDRECD4 | f2.0 | 3=Stay Away Cond, 4=House Arrest 24HRs, |
| 58-59 | SCDRECD5 | f2.0 | 5=Rpt to Prob upon Rel, 6=Rpt to Parole, 7=Halfway House, 8=Work-Rel from Jail, 9=High Risk Custody, 10=Medium Custody, 11=Stay Away Cond, 12=Rpt to Sent Judge, 13=Custodian, 17=Missing, 18=??????? |
| 60-61 | SAFPROB1 | f2.0 | Safety Problem-New Scheme |
| 62-63 | SAFPROB2 | f2.0 | Table of Outcomes for SAFPROB1-SAFPROB7 |
| 64-65 | SAFPROB3 | f2.0 | 1=D-Alcohol NT, 2=D-Drugs NT, 3=D Mental, |
| 66-67 | SAFPROB4 | f2.0 | 4=D + Prior D, 5=D + Prob D, 6=D + Pending D, |
| 68-69 | SAFPROB5 | f2.0 | 7=D + Parole D, 8=Anything + Prob D, |
| 70-71 | SAFPROB6 | f2.0 | 9=D And Prob No D, 10=D + Juvenile, |
| 72-73 | SAFPROB7 | f2.0 | 11=High Risk Vio, 12=Medium Risk Vio, 13=Threaten Witness, 14=Threaten Juror, 15=On Parole Danger, 16=D Plus Parole No D, 17=D + Weapon, 18=D Plus Alcohol Treatment, 19=D + Drug Treatment, 20=D Psych Treatment, 21=D + Prior D, 22=D + Unsup Prob, 23=D + Charge, 24=Anything + D, 25=Felony + Prior Juv, 26=Prior D Conviction, 27=Alive Witness, 28=Missing, 29=??????? |
| 74 | ALCHRECD | f1.0 | Alcohol Treatment Received-Safety 1=Enter Treatment, 2=Maintain Treatment, 3=Missing, 4=??????? |
| 75 | DRUGRECD | f1.0 | Drug Treatment Received-Safety 1=Enter Treatment, 2=Maintain Treatment, 3=Missing, 4=??????? |
| 76 | PSYRECD | f1.0 | Psychiatric Treatment Received-Safety 1=Competency Screening, 2=Maintain Treatment, 3=Missing, 4=??????? |
| 77 | CURFEWRD | f1.0 | Curfew Received-Safety 0=Blank Field, 1=Curfew Received, 2=Missing, 3=??????? |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 78-79 | FLTPROB1 | f2.0 | Flight Problem-New Scheme |
| 80-81 | FLTPROB2 | f2.0 | 1=Alcohol No Program, 2=Drug Use No Program, |
| 82-83 | FLTPROB3 | f2.0 | 3=Mental Obs, 4=Nonarea Resident, |
| 84-85 | FLTPROB4 | f2.0 | 5=Illegal Alien, 6=Identity Conflict, |
| 86-87 | FLTPROB5 | f2.0 | 7=No Fixed Address, 8=Prob Violation, |
| 88-89 | FLTPROB6 | f2.0 | 9=Prob Unsatisfied, 10=Parole Violation, |
| 90-91 | FLTPROB7 | f2.0 | 11=Parole Unsatisfied, 12=High Risk Violation, 13=Medium Risk Violation, 14=Warrant Outstanding, 15=Fugitive With FTA, 16=Present BRA, 17=2 Cases Pending, 18=Prior Sol Pros, 19=BRA Conviction, 20=AWOL, 21=Flee, 22=On Prob Non D, 23=On Parole Non D, 24=Under Sentence, 25=In Treatment-Alch, 26=In Treatment-Drugs, 27=Mental, 28=Non-Resid Verified, 29=Alien With Passport, 30=Conflict Address, 31=No Returnable Address, 32=Unsup Prob, 33=Unverified Address, 34=Defendant Ignorance, 35=Condition Violator, 36=Active Military, 37=Missing, 38=Blank, 39=??????? |
| 92 | ALCHRECA | f1.0 | Alcohol Treatment Received-Flight (same code as column 74) 1=Enter Treatment, 2=Maintain Treatment, 3=Missing, 4=??????? |
| 93 | DRUGRECA | f1.0 | Drug Treatment Received-Flight (same code as column 75) 1=Enter Treatment, 2=Maintain Treatment, 3=Missing, 4=??????? |
| 94 | PSYRECA | f1.0 | Mental Program Received-Flight (same code as column 76) 1=Competency Screening, 2=Maintain Treatment, 3=Missing, 4=??????? |
| 95-96 | OTHFLTR1 | f2.0 | Other Flight Condition Received |
| 97-98 | OTHFLTR2 | f2.0 | Table of Outcomes for OTHFLTR1-OTHFLTR4 |
| 99-100 | OTHFLTR3 | f2.0 | 1=Enroll in Alch Prog, 2=Enroll in Drug Prog, |
| 101-102 | OTHFLTR4 | f2.0 | 3=Interstate Superv, 4=Send Notice to..., 5=Live At, 6=Curfew, 7=Rpt to PSA in Person, 8=Rpt to Prob-Par upon Rel, 9=House Arrest, 10=Rpt to MPD Weekly, 11=Surrender Passport, 12=Stay in Alch Prog, 13=Maintain Drug Prog, 14=Competency Screening, 15=Stay Away Cond, 16=Live at, 17=Employment, 18=Student Status, 19=Rpt weekly to PSA, 20=Maint Psych Treat, 21=Other Condition, 22=Custody, 23=Judicial Order, 24=Mental Observ, 25=Attorney Conditions, 26=Person for Notice, 27=Stay Away from Place, 28=Rpt to Attorney, 29=Blank, 30=00's, 31=99's, 32=??????? |
| 103 | CITREL | f1.0 | Citation Release 1=Released, 2=Not Released, 3=Other |
| 104-105 | TLTAPPCD | f2.0 | Total Appearance Conditions Set |
| 106-107 | TLTSAFCD | f2.0 | Total Safety Conditions Set |
| 108-109 | TLTCDS | f2.0 | Total Conditions Set |
| 110 | RELTOI1 | f1.0 | Type of Interview at Release 1 1=C-L, 2=Lock-Up, 3=GJO, 4=Other, 5=Citation |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 111-112 | RELEASE | f2.0 | Initial Court Action 1=PR, 2=PR With Conds, 3=Percentage, 4=Percent With Conds, 5=Cash Bond, 6=Cash With Conds, 7=Cash-Surety Option, 8=Cash-Surety Conds, 9=Surety Bond, 10=Surety With Conds, 11=Prev Det Hearing, 12=5-Day Hold, 13=Hold WO Bond, 14=No Paper, 15=Dismissed, 16=Competency Screening, 17=GJ Original, 18=Indictment, 19=Plea, 20=Fugitive Returns, 21=Station House Bond, 22=UAB, 23=UAB With Conds, 24=Diversion, 25=Unknown, 26=Mental Observation, 27=Work Release, 28=____, 29=____, 30=Turned over to..., 31=Missing, 32=Community Services, 33=____, 34=Blank Field, 35=?????? |
| 113 | RELCT1 | f1.0 | Court of Initial Action 1=Superior CT, 2=US Magistrate, 3=US District CT |
| 114-117 | RELJUD1 | f4.0 | Initial Release Judge |
| 118-119 | RPTSET1 | f2.0 | Report Condition Set-Old Scheme 1=Yes, 2=Missing, 3=???? |
| 120 | RPTYPE1 | f1.0 | How to Report 1=By Phone, 2=In Person |
| 121-126 | BONDAMT | f6.0 | Bond Amount Set |
| | | | Actual Bond Posting Date |
| 127-128 | BDPSTYR | f2.0 | Year |
| 129-130 | BDPSTMO | f2.0 | Month |
| 131-132 | BDPSTDY | f2.0 | Day |
| 133 | BONDPOST | f1.0 | Bond Posted? 1=Posted, 2=Not Posted |
| 134-137 | POSTIME | f4.0 | Time to Posting |
| 138-139 | APPCD1 | f2.0 | Appearance Conditions |
| 140-141 | APPCD2 | f2.0 | Table of Outcomes for APPCD1-APPCD5 |
| 142-143 | APPCD3 | f2.0 | 1=Enroll in Alch Program, |
| 144-145 | APPCD4 | f2.0 | 2=Enroll in Drug Prog, 3=Interstate Superv, |
| 146-147 | APPCD5 | f2.0 | 4=Send Notice To..., 5=Live At, 6=3rd Party-Person, 7=Provide PSA Address, 8=Report Weekly, 9=Report Prob-Parole-Judg, 10=Custody, 11=Halfway House, 12=Work Rel from Jail, 13=Surrender Passport, 14=Stay in Alch Prog, 15=Stay in Drug Prog, 16=Competency Screening, 17=Rpt to Armed Forces, 18=Stay in Area, 19=Maint Psych Treat, 20=Complaining Witness, 21=Post-Rel Interview, 22=Other Reporting, 23=No Rearrest, 24=Other Cond, 25=Curfew 26=Seek or Keep Job, 27=Stay-Enter School, 28=Stay in Area, 29=Judicial Order, 30=24hr Resident Custod, 31=Pay Attorney, 32=Custody Halfway Hse, 33=Trial Priority, 34=Held WO Bond, 35=Mental Observation, 36=Missing, 37=???? |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 148-149 | SAFCD1 | f2.0 | Safety Condition Set (Actual CT Action) |
| 150-151 | SAFCD2 | f2.0 | Table of Outcomes for SAFCD1-SAFCD5 |
| 152-153 | SAFCD3 | f2.0 | 1=Rpt to Armed Forces, 2=Notify PSA of Addres, 3=Resident Custody, |
| 154-155 | SAFCD4 | f2.0 | 4=Stay Away Order, 5=Report Weekly, |
| 156-157 | SAFCD5 | f2.0 | 6=Rpt Prob or Parole, 7= Custody Corrections, 8=WorkRel, 9=Custody, 10=Stay in Area, 11=Curfew, 12=No Rearrest, 13=Other Cond, 14=Reside At Cond, 15=Employment Cond, 16=Student Status, 17=Judicial Order, 18=Pay Attorney, 19=Missing, 20=??????? |
| 158-159 | CUSTODY | f2.0 | Third Party Custody Program (CT Ordered) 1=BonaBond, 2=Bureau, 3=CIRO, 4=Halfway Hse, 5=Other, 6=Person, 7=RAP, Inc, 8=RCA, 9=Stepping Stones, 10=Suitable, 11=Missing, 12=Dept of Corr, 13=Military Police, 14=Project Triangle, 15=AYUDA, 16=Comm Reality Proj, 17=Blackman's, 18=St Elizabeth's, 19=????? |
| 160-161 | DETAPP1 | f2.0 | Detailed Appearance Condition Set (CT Ordered) |
| 162-163 | DETAPP2 | f2.0 | Table of Outcomes for DETAPP1-DETAPP5 |
| 164-165 | DETAPP3 | f2.0 | 0=Blank, 1=Enroll in Alch Prog, |
| 166-167 | DETAPP4 | f2.0 | 2=Enroll in Drug Prog, 3=Interstate Superv, |
| 168-169 | DETAPP5 | f2.0 | 4=Send Notice To..., 5=Live At, 6=Curfew, 7=Rpt to PSA in Person, 8=Rpt to Prob-Par Upon, 9=House Arrest, 10=Rpt to MPD Weekly, 11=Surrender Passport, 12=Stay in Alch Prog, 13=Maintain Drug Prog, 14=Competency Screening, 15=Stay Away CW, 16=Live At, 17=Employment, 18=Student Status, 19=Rpt Weekly to PSA, 20=Maint Psych Treat, 21=Other Cond, 22=Custody, 23=Judicial Order, 24=Mental Obser, 25=Attorney Conditions, 26=Person for Notice, 27=Stay Away frm Place, 28=Rpt to Attorney, 29=Missing, 30=?????? |
| 170-171 | DETSAF1 | f2.0 | Detailed Safety Conditions Set (CT Ordered) |
| 172-173 | DETSAF2 | f2.0 | Table of Outcomes for DETSAF1-DETSAF3 |
| 174-175 | DETSAF3 | f2.0 | 0=Blank, 1=Enroll in Alch Prog, 2=Enroll in Drug Prog, 3=Stay Away Cond, 4=Rpt to Prob-Par, 5=Send Notice, 6=Stay in Alch Prog, 7=Stay in Drug Prog, 8=Competency Screening, 9=Sty Away frm Place, 10=Curfew, 11=Stay in Psych Prog, 12=Other Cond, 13=Rpt Weekly, 14=Seek-Keep Job, 15=Surrender Passport, 16=Address Cond, 17=Judicial Order, 18=Handwriting Sample, 19=Mental Obs, 20=Interstate Superv, 21=Pay Attorney, 22=Student Status, 23=Call Attorney Weekly, 24=Missing, 25=?????? |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 176 | STAYAT | f1.0 | Stay At Condition-New Scheme 1=High Risk Curfew, 2=Medium Risk Curfew, 3=Missing |
| 177 | D1 | f1.0 | First Part of Docket-Misd-Felony Status 1=Sup CT-Felony, 2=Sup CT-Misd, 3=Missing, 4=Blank, 5=District CT |
| 178 | ATTSTAT | f1.0 | Attorney Status 1=CJA-100%,2=CJA-Less 3=No Lawyer Appointed, 4=PDS, 5=Retained, 6=Student, 7=Unknown, 8=Missing, 9=???? |
| 179 | CODEFT | f1.0 | Codefendant? 1=Yes, 2=No, 3=Missing, 4=???? |
| 180-182 | CHARGE1 | f3.0 | Most Serious Charge At Arrest |
| 183-185 | CHARGE2 | f3.0 | Second Most Serious Charge At Arrest |
| 186-191 | JDGDATE | f6.0 | Judgment Date |
| | | | Bench Warrant Issuance Date |
| 192-193 | FTAYR1 | f2.0 | Year |
| 194-195 | FTAMN1 | f2.0 | Month |
| 196-197 | FTADY1 | f2.0 | Day |
| 198 | FTAREA1 | f1.0 | Court's Reason for Bench Warrant (Above) 1=FTA-PR, 2=FTA Cash Bond, 3=Failure to Pay Fine, 4=Other, 5=FTA-Surety Bond, 6=Vio Of Court Order, 7=Missing, 8=Probation Violation, 9=?????? |
| | | | Bench Warrant Clearance Date |
| 199-200 | FTADYR1 | f2.0 | Year |
| 201-202 | FTADMN1 | f2.0 | Month |
| 203-204 | FTADDY1 | f2.0 | Day |
| 205 | HOWDISP1 | f1.0 | How Was Bench Warrant Cleared? 1=Quashed, 2=Executed, 3=Expired, 4=Missing, 5=Blank Field, 6=?????? |
| 206 | BRA1 | f1.0 | Defendant Charged for FTA 0=Blank, 1=Yes, 2=No, 3=Missing, 4=Unknown, 5=???? |
| 207 | BWCOUNT | f1.0 | Number of Bench Warrants Issued |
| 208-209 | CTDATES | f2.0 | Total Court Dates Scheduled (Missed + Made) |
| 210-211 | TLTNOT | f2.0 | Total Appearances Notified by PSA |
| 212-213 | TLTACK | f2.0 | Total Notices Acknowledged to PSA |
| 214-215 | CLFTC | f2.0 | Number of Violated Conditions |
| 216 | VIOHEAR | f1.0 | Violation Hearing Held? 1=Yes, 2=No |
| 217 | RMAILCT | f1.0 | Return Mail Count |
| 218-219 | RMAILRES | f2.0 | Return Mail Reason |
| 220 | FTCRES1 | f1.0 | Failure to Comply-Surrender Passport 1=Yes, 2=No |
| 221 | FTCRES2 | f1.0 | FTC-Curfew-House Arrest 1=Yes, 2=No |
| 222 | FTCRES3 | f1.0 | FTC-Custody Program or Person 1=Yes, 2=No |
| 223 | FTCRES4 | f1.0 | FTC-Complaining Witness 1=Yes, 2=No |
| 224 | FTCRES5 | f1.0 | FTC-Residence-No Address-Area 1=Yes, 2=No |
| 225 | FTCRES6 | f1.0 | FTC-Reporting Condition 1=Yes, 2=No |
| 226 | FTCRES7 | f1.0 | FTC-Drug Program Condition 1=Yes, 2=No |
| 227 | FTCRES8 | f1.0 | FTC-Alch Program Condition 1=Yes, 2=No |
| 228 | FTCRES9 | f1.0 | FTC-Interstate Supervision 1=Yes, 2=No |
| 229-236 | PSAID | f8.0 | Pretrial Services Agency I.D. Number |
| 237-239 | PDID1 | f3.0 | Police Identification Number |
| 240-242 | PDID2 | f3.0 | |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 243-248 | DOBDATE | f6.0 | Date Of Birth |
| 249-250 | BIRTHPL | f2.0 | Birth Place |
| 251-252 | AGE | f2.0 | Age of Defendant at time of Court |
| 253 | SEX | f1.0 | Sex 1=male, 2=female |
| 254 | RACE | f1.0 | Race 1=black, 2=white, 3=other |
| 255-256 | MARRY | f2.0 | Marital Status |
| 257 | LWS | f1.0 | Lives with Spouse 1=yes, 2=no |
| 258 | LWC | f1.0 | Lives with Children 1=yes, 2=no |
| 259-260 | KIDS | f2.0 | Number of Children |
| 261-263 | EDUCATE | f3.0 | Years in School |
| 264 | FAMCT | f1.0 | Number of Family in Area |
| 265 | ALIEN | f1.0 | Alien 1=yes, 2=no |
| 266 | QUAD | f1.0 | Area of City 1=NW, 2=NE, 3=SE, 4=SW |
| 267-268 | STATE | f2.0 | State of Residence |
| 269-273 | ZIPCODE | f5.0 | Zipcode of Residence |
| 274 | CANRET | f1.0 | Can Return To Previous Residence 1=yes |
| 275 | HSETYP | f1.0 | Type Of Residence, 1=house,2=room, 3=apartment,4=hotel,5=employer,6=motel |
| 276 | BUYRENT | f1.0 | Buying? 1=Buying,2=renting,3=neither |
| 277 | LIVECW | f1.0 | Lives With Complaining Witness 1=yes |
| 278-279 | LIVEWITH | f2.0 | Person Living With Accused |
| 280 | EMPSTAT | f1.0 | Employment Status,1=Employed, 2=Unemployed,3=Homemaker,4=Other |
| 281 | STUDENT | f1.0 | In School, 1=yes, 2=no |
| 282 | WORKSTAT | f1.0 | Fully Employed? 1=Full Time,2=Partime |
| 283-286 | PAYAMT | A4 | Amount of Pay |
| 287-288 | HOURLY | f2.0 | Frequency Of PAYAMT,1=Bi-Week,2=Hourly |
| 289 | EMOTPROB | f1.0 | Emotional Problem, 1=yes,2=no |
| 290 | EMOTSTAT | f1.0 | Status of EMOTPROB, 1=current,2=prior |
| 291 | PHYSPROB | f1.0 | Physical Problem, 1=yes, 2=no |
| 292 | PHYSTAT | f1.0 | Status of PHYSPROB, 1=current,2=prior |
| 293 | DRUGPROB | f1.0 | Drug Problem, 1=yes, 2=no |
| 294-295 | DRUGTYPE | f2.0 | Type of Drug |
| 296 | DRUGSTAT | f1.0 | Status of Drug Problem 1=current, 2=prior |
| 297 | ALCHPROB | f1.0 | Alcohol Problem 1=yes, 2=no |
| 298 | ALCHSTAT | f1.0 | Status of Alcohol Problem 1=current 2=prior |
| 299 | PROBATN | f1.0 | On Probation 1=yes, 2=no |
| 300-301 | PROBADJ | f2.0 | Probation Adjustment 1=good, 2=marginal,3=poor,4=satisfactory |
| 302 | PAROLE | f1.0 | On Parole 1=yes, 2=no |
| 303-304 | PARADJ | f2.0 | Parole Adjustment 1=good, 2=marginal, 3=poor, 4=satisfactory |
| 305 | PROPARCT | f1.0 | Count of Probation-Parole Status |
| 306-307 | EXCON | f2.0 | Count of Prior Convictions |
| 308-313 | CV1 | f6.0 | First Conviction Date |
| 314-319 | CV2 | f6.0 | Most Recent Conviction Date |
| 320-321 | PENDCASE | f2.0 | Number of Pending Cases |
| 322-327 | PENDDATE | f6.0 | Date of Pending Case |
| 328-329 | REARREST | f2.0 | Number of Rearrest Cases Before Disposition |
| 330-335 | READATE | f6.0 | Date of First Rearrest |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 336-337 | COURTCT | f2.0 | Number of Court Check Ins To PSA |
| 338-339 | RPTCOUNT | f2.0 | Number of Reports To PSA |
| 340-341 | OTHERCT | f2.0 | Other Check-Ins To PSA |
| 342 | STATUS | f1.0 | Disposition Status, 1=closed,2=open |
| 343-344 | OUTCOME | f2.0 | Final Disposition of Case |
| 345-346 | CHARGEN | f2.0 | Most Serious Charge At Arrest 2=rape,3=burglary |
| 347 | CHGGP | f1.0 | 4=drugs,6=flight,7=forgery,8=fraud,10=murder 11=kidnap,12=larceny,13=robbery,15=prostitution 16=auto theft,17=stolen property,18=weapons 19=possession implements of crime,20=destruction of property |
| 348-349 | CHARGEDV | f2.0 | Most Serious Charge For Dangerous And Violent Crimes |

THE NEXT FOUR VARIABLES SUMMARIZE THE PREARREST, PRETRIAL, AND POSTDISPOSITION ARREST EXPERIENCE FOR THIS ARRESTED INDIVIDUAL

| | | | |
|---|---|---|---|
| 350-351 | ARRSTNO | f2.0 | Number of arrests Jan 1980-Dec 1982 |
| 352-353 | PREARR | f2.0 | Number of prearrest arrests |
| 354-355 | PREDISP | f2.0 | Number of pre-disposition arrests |
| 356-357 | POSTDISP | f2.0 | Number of post-disposition arrests |

VARIABLES TAKEN FROM THE PREARREST DATA RECORD OR RECORDS AND GIVING DETAILS ABOUT THE ACCUSED AT TIME OF THE TWO MOST RECENT PREARREST ARRESTS

| | | | |
|---|---|---|---|
| 358-363 | PARRDAT1 | f6.0 | Date of prearrest arrest #1 |
| 364-365 | PARRREL1 | F2.0 | Release variable for prearrest #1, see 111-112 |
| 366-368 | PARRCH1 | F3.0 | Charge 1 for prearrest #1, see 180-182 |
| 369 | PARRBW1 | F1.0 | Bench warrant dummy for prearrest #1 see #207 |
| 370 | PARRPRO1 | F1.0 | Probation dummy variable for prearrest #1, see #299 |
| 371 | PARRPAR1 | F1.0 | Probation dummy variable for prearrest #1, see #302 |
| 372-373 | PARREXC1 | F2.0 | Exconvict dummy for prearrest #1, see 306-7 |
| 374-375 | PARRCHN1 | F2.0 | CHARGEN code for prearrest #1,see 345-6 |
| 376-377 | PARRCHV1 | F2.0 | CHARGEDV code for prearrest #1, see 348-49 |
| 378-383 | PARRDAT2 | F6.0 | Date of prearrest arrest #2, note that this is blank if there is only one prearrest. If there are two or more prearrests, this is the most recent prearrest arrest. |
| 384-385 | PARRREL2 | F2.0 | Release variable for prearrest arrest #2 |
| 386-388 | PARRCH2 | F3.0 | Charge 1 for prearrest arrest #2 |
| 389 | PARRBW2 | F1.0 | Bench warrant for prearrest #2 |
| 390 | PARRPRO2 | F1.0 | Probation |
| 391 | PARRPAR2 | F1.0 | Parole |
| 392--393 | PARREXC2 | F2.0 | Exconvict |
| 394-395 | PARRCHN2 | F2.0 | Chargn for prearrest arrest #2 |
| 396-397 | PARRCHV2 | F2.0 | Charge Dangerous and Violent for Prearrest #2 |

THESE VARIABLES ARE TAKEN FROM THE PRETRIAL OR PREDISPOSITION ARREST RECORDS AND GIVE DETAILS ABOUT THE ACCUSED AT THE TIME OF PRETRIAL ARREST

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 398-403 | PDISDAT1 | F6.0 | Date of predisposition arrest #1. This is the predisposition arrest closest to the arrest |
| 404-405 | PDISREL1 | F2.0 | Release variable for predisp arrest #1 |
| 406-408 | PDISCH1 | F3.0 | Charge1 for predisposition arrest #1 |
| 409 | PDISBW1 | F1.0 | Bench warrant for predisposition arrest #1 |
| 410 | PDISPRO1 | F1.0 | Probation |
| 411 | PDISPAR1 | F1.0 | Parole |
| 412-413 | PDISEXC1 | F2.0 | Number of Prior Convictions |
| 414-415 | PDISCHN1 | F2.0 | Chargen for predisposition arrest #1 |
| 416-417 | PDISCHV1 | F2.0 | Charge dangerous or violent for predisposition arrest #1 |
| 418-423 | PDISDAT2 | F2.0 | Date of predisposition arrest #1 |
| 424-425 | PDISREL2 | F2.0 | Release variable for predisp arrest #2 |
| 426-428 | PDISCH2 | F2.0 | Charge 1 for predisposition arrest #2 |
| 429 | PDISBW2 | F1.0 | Bench warrant for predisposition arrest #2 |
| 430 | PDISPRO2 | F1.0 | On probation |
| 431 | PDISPAR2 | F1.0 | On parole |
| 432-433 | PDISEXC2 | F2.0 | Number of prior convictions |
| 434-435 | PDISCHN2 | F2.0 | Chargen for predisposition arrest #2 |
| 436-437 | PDISCHV2 | F1.0 | Charge dangerous or violent for predisposition arrest #2 |
| 438-443 | PDISDAT3 | F6.0 | Arrest date for third most recent predisposition arrest in days counted from 1900 using SAS time measure |
| 444-445 | PDISREL3 | F2.0 | Release variable for 3rd predisposition arrest |
| 446-448 | PDISCH3 | F2.0 | Most serious charge for predisposition arrest #3 |
| 449 | PDISBW3 | F1.0 | Bench warrant issued after arrest #3 |
| 450 | PDISPRO3 | F1.0 | On probation at time of predisposition arrest #3 |
| 451 | PDISPAR3 | F1.0 | On parole at time of predisposition arrest #3 |
| 452-453 | PDISEXC3 | F2.0 | Number of prior convictions at arrest #3 |
| 454-455 | PDISCHN3 | F2.0 | Most serious charge at arrest #3 (CHARGN) |
| 456-457 | PDISCHV3 | F2.0 | Dangerous or violent charge at arrest #3 |

VARIABLES TAKEN FROM THE POSTDISPOSITION ARREST RECORDS OF THE ACCUSED GIVING DETAILS ABOUT THE ACCUSED AT THE TIME OF THE POSTDISPOSITION ARRESTS

| | | | |
|---|---|---|---|
| 458-463 | POSTDAT1 | F6.0 | Arrest date for first postdisposition arrest in days using SAS time counter variable |
| 464-465 | POSTREL1 | F2.0 | Release variable for postdisposition arrest #1 |
| 466-468 | POSTCH1 | F2.0 | Most serious charge,postdisposition arrest #1 |
| 469 | POSTBW1 | F2.0 | Bench warrant issued after arrest #1 |
| 470 | POSTPRO1 | F1.0 | On probation at time of arrest #1 |
| 471 | POSTPAR1 | F1.0 | On parole at time of arrest #1 |
| 472-473 | POSTEXC1 | F2.0 | Number of prior convictions at arrest #1 |
| 474-475 | POSTCHN1 | F2.0 | Chargn for postdisposition arrest #1 |
| 476-477 | POSTCHV1 | F2.0 | Chargedv for postdisposition arrest #1 |
| 478-483 | POSTDAT2 | F6.0 | Arrest date for second post disposition arrest in days using SAS time variable |
| 484-485 | POSTREL2 | F2.0 | Release for postdisposition arrest #2 |
| 486-488 | POSTCH2 | F2.0 | Most serious charge at arrest for arrest #2 |

| COLUMN | VARIABLE | FORMAT | DESCRIPTION |
|---|---|---|---|
| 489 | POSTBW2 | F1.0 | Bench warrant issued for arrest #2 |
| 490 | POSTPRO2 | F1.0 | On probation at postdisposition arrest #2 |
| 491 | POSTPAR2 | F1.0 | On parole at postdisposition arrest #2 |
| 492-493 | POSTEXC1 | F2.0 | Number of prior convictions at arrest #2 |
| 494-495 | POSTCHN1 | F2.0 | Chargn for postdisposition arrest #2 |
| 496-497 | POSTCHV1 | F2.0 | Chargedv for postdisposition arrest #2 |

APPENDIX D

BIBLIOGRAPHY

Abowd, John M. and Henry S. Farber, [1982] "Job Queues and the Union Status of Workers," Industrial and Labor Relations Review, Vol. 35, pp 354-367.

Amemiya, Takeshi, [1981] "Qualitative Response Models: A Survey," Journal of Economic Literature, Vol. XIX, No. 4, pp 1483-1536.

Becker, Gary and William Landes, [1974] Essays in the Economics of Crime and Punishment (New York: National Bureau of Economic Research)

Berndt, Earnest R., Bronwin H. Hall, Robert E. Hall, and Jerome A. Hausman, [1974] "Estimation and Inference in Non-Linear Structural Models," Annals of Economic and Social Measurement Vol. 3, (October) pp 653-665.

Bureau of Justice Statistics, [1985], "Pretrial Release and Misconduct," Special Report (January).

Clark, C., [1961] "The Greatest of a Finite Set of Random Variables," Operations Research, Vol. 9, pp 145-162.

Daganzo, C., F. Bouthelier, and Y. Sheffi, [1977] "Multinomial Probit and Qualitative Choice - A Computationally Efficient Algorithm, Transportation Science, Vol. 11, pp 338-358.

Farber, Henry S. [1983] "Worker Preference For Union Representation," Research in Labor Economics Supplement Vol. 2, pp 171-205.

Fishe, Raymond H.P., Robert P. Trost, and Philip Lurie, [1981] "Labor Force Earnings and College Choice Among Young Women: An Examination of Selectivity Bias and Comparative Advantage," Economics of Education Review pp 169-191.

Goldberg, I. and F.C. Nold, [1980] "Does Reporting Deter Burglars? - An Empirical Analysis of Risk and Return to Crime," Review of Economics and Statistics (August) pp 424-431.

Goldkamp, John S., Michael R. Gottfredson, and Susan Michell-Herzfeld, [1981] "Bail Decisionmaking: A Study in Policy Guidelines," Report to the National Insititute of Corrections, U.S. Department of Justice.

Goldkamp, John S., [1983] "Questioning the Practice of Pretrial Detention: Some Empirical Evidence from Philadelphia," The Journal of Criminal Law and Criminology Vol. 74 No. 4, pp 116-148.

Hausman, Jerome and David Wise, [1978] "A Conditional Probit Model for Qualitative Choice" Econometrica Vol. 46, pp 403-426.

Heckman, James K., [1979], "Sample Selection Bias as a Specification Error," Econometrica, Vol. 47 No. 1, pp 153-161.

Lee, Lung-Fei, [1984] "Sequential Discrete Choice Econometric Models With Selectivity," Mimeograph, University of Minnesota.

Lee, Lung-Fei and G.S. Maddala, [1983] "Sequential Decision Rules and Selectivity in Discrete Choice Econometric Models," Mimeograph, Center for Econometrics and Decision Sciences, University of Florida.

Lerman, Steven R. and Charles F. Manski, [1981] "On the Use of Simulated Frequencies to Approximate Choice Probabilities," in Structural Analysis of Discrete Data with Econometric Applications edited by Charles Manski and Daniel McFadden (Cambridge, Massachusetts: MIT Press), pp 305-319.

Mackaay, Ejan, [1982] Economics of Information and Law, (Boston: Kluwer-Nijhoff).

Maddala, G.S., [1983] Limited Dependent and Qualitative Dependent Variables in Econometrics, (Cambridge: Cambridge University Press)

McFadden, Daniel, [1975] "Conditional Logit Analysis of Qualitative Choice Behavior," in Frontiers in Econometrics edited by Paul Zarembka, (New York: Academic Press), pp 105-142.

Meng, Chung-Lo and Peter Schmidt, [1985] "On the Cost of Partial Observability in the Bivariate Probit Model," International Economic Review, Vol. 26 No. 1, (February) pp 71-85.

Myers, Samuel, [1981] "The Economics of Bail Jumping," Jounal of Legal Studies (June).

Porier, Dale J., [1980] "Partial Observability in the Bivariate Probit Model," Journal of Econometrics, Vol. 44, (February) pp 210-217.

Posner, Richard A., [1975] "The Economic Approach to Law," Texan Law Review pp 757-782

Rhodes, William, et. al. [1984] "Pretrial Release and Misconduct in Federal District Courts," Report to the Bureau of Justice Statistics by INSLAW, Inc., (December).

Rhodes, William, [1985] "The Adequacy of Statistically Derived Prediction Instruments in the Face of Sample Selectivity: Criminal Justice as an Example," Mimeograph, INSLAW, Inc.

Roth, Jeffrey A., and Paul B. Wice, [1980] Pretrial Release and Misconduct in the District of Columbia PROMIS Research Project Publication 16, (April).

Sherwood-Fabre, Liese, [1984] "An Experiment in Bail Reform: Evaluating Pretrial Release Services Agencies in Federal District Courts," Report to the National Institute of Justice, U.S. Department of Justice.

Toborg, Mary, et al. [1984] "Pretrial Release Assessment of Danger and Flight," Report to the District of Columbia Pretrial Services Agency sponsored by the National Institute of Justice.

Trost, Robert P. and Anthony M.J. Yezer, [1985] "Sequential Selection and Selectivity in a Model of the Market for Bail Bond," Business and Economics Section, Proceedings of the American Statistical Association.

Trost, Robert P. and Anthony M.J. Yezer, [1986] "Crooks and Their Keepers: Predicting and Preventing Criminal Behavior," paper presented at the Conference on Asymmetric Information Models in Economics, University of California, Santa Cruz (May).

Witte, Ann D. and Peter Schmidt, [1979] "An Analysis of the Type of Criminal Activity Using the Logit Model," Journal of Research in Crime and Delinquency, (January) pp 164-179.

Zellner, Arnold and T.H. Lee, [1965] "Joint Estimation of Relationships Involving Discrete Random Variables," Econometrica, Vol. 33 (May) pp 382-394.