126663

The Termination Rate of Adult Criminal Careers*

Alfred Blumstein
Jacqueline Cohen
Andrew Golub

December 20, 1989

*This research was supported in part by grant 86-IJ-CX-0047 from the National Institute of Justice, U. S. Department of Justice. The views expressed are those of the authors and do not necessarily reflect the opinions of the funding agency.

The authors also acknowledge and thank John Lehoczky, Mark Schervish, and Steve Garber for their suggestions during the developmental phases of this research.

ABSTRACT

The adult arrest histories of male criminal offenders arrested in the Detroit SMSA between January 1974 and December 1977 are studied to determine the rate at which offenders terminate criminal activity and how this rate varies based on offenders' prior criminal record. Arrestees are first divided into groups on the basis of similar demographic information and criminal records prior to the time of a chosen arrest. Average termination rates are estimated for each group based on criminal activity after the chosen arrest using maximum likelihood techniques.

The primary variations in δ observed in this analysis are that termination rates generally decrease as the number of prior arrests increase for white, 17-29 year-old, offenders (black offenders did not exhibit this variation), and that black and white, 30-39 year-old offenders have a lower termination rate than offenders who are 17-29 or over 40 years old at the time of arrest.

TABLE OF CONTENTS

1.0 INTRODUCTION

1.1 The Criminal Careers Approach

The criminal careers approach focuses on the longitudinal analysis of individual offenders' criminal activity. This can be contrasted with the analysis of aggregate crime statistics which focuses on the total number of arrests for a population at a specific time. A National Research Council panel on research on criminal careers identifies four key dimensions which characterize criminal careers:[1]

1. Participation - The distinction between those who engage in crime and those who do not.

2. Frequency (μ) - The rate of criminal activity of those who are active.

3. Seriousness of offenses committed.

4. Career Length - The length of time an offender is active.

The termination rate (δ) -- defined as the probability an offender terminates criminal activity in a given year -- is an alternative way to characterize the distribution of career lengths. As an example, if an offender's termination rate is constant throughout his criminal career, $\delta = .10$ per year, then the expected career length is $1/\delta = 10$ years.[2] In this respect career length and termination rate are inversely related.

---

[1] Blumstein, Cohen, Roth and Visher, 1986, pp. 1-5.

[2] Since in this example, the termination rate is assumed to be constant, the distribution of possible career lengths is exponentially distributed, and the mean of an exponential distribution is equal to the inverse of the termination rate.

1.2 The Value of Termination Rate Information

The focus of this study is estimation of the termination rate from offender follow-up arrest reports, and an analysis of how it varies across offender attributes. Knowledge of termination rates can be useful in the following types of criminal justice policy decisions:

1. Career length information, as one of the characteristics of criminal careers, is useful for projecting the demand for criminal justice facilities and services. For example, Barnett (1987) used the career termination rate (along with other information) to project the size of the future prison populations in Massachusetts, Florida, and Utah. His concern centered around whether additional prison facilities need to be constructed or not.

2. Career length information can be used as one of the criterion for establishing sentencing guidelines in order to maximize crimes averted due to incapacitation.[3] Once an offender's criminal career has ended, incarceration no longer serves to avert criminal activity by that offender. Therefore, career length information is a necessary part of the estimate of the expected number of crimes averted through incapacitation by alternative sentencing policies.

3. The identification of exogenous individual and societal factors that influence criminal career length can be used to determine where public policies should be directed in order to decrease crime. For example, Blumstein, Cohen and Hsieh (1982) found evidence that among adult offenders, those in their twenties had a higher termination rate than offenders who were still criminally active in their thirties. This result implies that an effective crime reduction program might focus on discouraging twenty year-olds from ever starting criminal activity, while targeting efforts to end criminal careers at criminal offenders in their thirties.

---

[3]Due to current prison overcrowding, efficient use of the limited supply of prison space has received considerable attention. However other considerations, such as ethical constraints on the amount of deserved punishment, also play a central role in establishing sentencing guidelines.

1.3 Prior Research on Recidivism

Recidivism is a widely-used, traditional measure of individual follow-up offending. In this section the relationship between recidivism and criminal careers is explored. Furthermore, offender attributes which prior studies found to be related to recidivism are identified.

1.3.1 Recidivism and Criminal Careers

Criminal recidivism is defined as the future recurrence of criminal behavior by previously identified offenders. Maltz (1984) characterizes recidivism as a two component process, the probability that an individual offender will eventually recidivate, γ, and the distribution of recidivism times for those who do recidivate. An act of recidivism can be defined as either committing, being arrested, being found guilty, or being incarcerated for a criminal act.[4] In his most basic model Maltz assumes that the time to recidivism for recidivists is exponentially distributed with rate, Φ.

The recidivism model parameters (γ,Φ) are directly observable from follow-up data of offender criminal records. However, these parameters are not direct indicators of the distinct offender behavioral characteristics that comprise the various aspects of individual criminal careers. For a group of arrestees, the

---

[4]Of course, the rate of recidivism will vary with the definition that is used; the further the penetration into the criminal justice system that is required, the lower the recidivism rate.

proportion of offenders who do recidivate, γ, (whether defined as criminal offense, arrest, conviction or incarceration) confounds the frequency rate at which recidivist events occur and the likelihood of remaining criminally active. Only those offenders who remain active long enough to incur a recidivist event will be counted among eventual recidivists. Thus γ is a function of the rate at which offenders terminate, δ, and the rate at which re-arrest occurs, μ. Likewise, the rate parameter for times to recidivism, Φ, also confounds both of these aspects of criminal careers. Times to recidivism will be short for high frequency offenders who recidivate soon after release but will also be short for offenders who are highly prone to terminating criminal activity, since these offenders must either recidivate soon or not at all.

While recidivism reflects a combination of various criminal career parameters, it does not provide a direct measure of the termination rate, δ. The exact relationship between the recidivism parameters and the behavioral aspects of criminal careers will depend on the specific model that is invoked. For example, using a basic criminal career model (μ and δ time-invariant over an offender's lifetime), the parameters of Maltz's recidivism model (with recidivism times exponentially distributed with parameter Φ) are as follows: $\gamma = \mu / (\mu + \delta)$; and $\Phi = (\mu + \delta)$.[5] The relationships from recidivism to criminal careers parameters are

[5]See Appendix C for derivations.

as follows: $\delta = (1-\gamma)/\Phi$; $\mu = \gamma/\Phi$. These relationships between recidivism parameters and criminal careers parameters illustrate the dependence of each recidivism parameter, γ and Φ, on both the frequency of re-arrest and the rate at which offenders terminate offending, μ and δ, respectively.

1.3.2 Covariates of Recidivism

A number of studies have explored the covariates of recidivism. Schmidt and Witte (1989), using recidivism models similar to Maltz's, studied criminal offenders released from North Carolina prisons between July 1977 and June 1978, and between July 1979 and June 1980. The recidivism parameters were allowed to vary with offender attributes in order to detect covariates of recidivism. They found that the offenders more likely to return to prison and to return sooner were younger, black males with many prior incarcerations, who had drug or alcohol addictions, and whose prior incarceration was lengthy and was for a property offense.

Beck and Shipley (1989) analyzed recidivism of over 16,000 inmates released from prisons in the U.S. during 1983. They report that 62.5% of these offenders were re-arrested for a felony or serious misdemeanor within 3 years of release. Higher recidivism rates (proportion of offenders re-arrested) were exhibited by males, blacks, high school dropouts, those who were younger at release, those with more prior adult arrests, those currently incarcerated for property offenses, and those releasees who were younger when first arrested. In a multivariate analysis, re-arrest rates were estimated to be highest at 91% within three years after release for inmates who were under age 25 at release, had 7 or more prior adult arrests, had a prior escape or revocation, and who were released after serving time for a property offense. The lowest re-arrest rate was 17% within a three year follow-up estimated for inmates who were 35 or older at release, had 3 or fewer prior adult arrests, no escape or revocation record, and who were not serving time for a property offense.

These results on the relationship of prior record, age at release, and offense type to recidivism mirror similar findings previously reported in studies of other inmates (e.g., Hoffman and Beck, 1980; Greenwood, 1982; Rhodes et. al. 1983). These covariates of recidivism represent natural candidates for exploration as offender attributes that are related to termination rates.

1.4 Prior Research on Termination Rates

While the research literature on participation, frequency, and seriousness of criminal activity is quite large[6], relatively few analyses have examined criminal career length or termination rates. This lack of analysis is probably due to the difficulty inherent in trying to estimate with any precision the date when an individual actually terminates criminal activity.

One method of estimating the termination rate within a population is the life-table approach which uses the age distribution of individuals arrested in a given year to infer the age-specific termination rate.[7] Blumstein, Cohen, and Hsieh (1982) used life-table techniques to estimate the termination rates for adult arrestees in Washington, D.C., for the years 1970 through 1976 (each year was analyzed separately). Figure 1 shows the distribution of arrestees by age for 1973. This raw count of arrestees by age was adjusted for the following:

1. variations in the size of the base population at each age,
2. offenders' age at the start of adult criminal activity,
3. variations in participation and termination rates over time,
4. variation in arrest rates across age.

Once these various other factors that might influence the age

---

[6]Appendices A and B of Blumstein, et. al., 1986, contain an extensive review of this literature.

[7]This approach for estimating the criminal career termination rate was first suggested by Shinnar and Shinnar (1975).

distribution independently of career termination were accounted for, the remaining changes in the size of the adjusted population by age and particularly the decline in arrestees with age, imply age-specific termination rates.

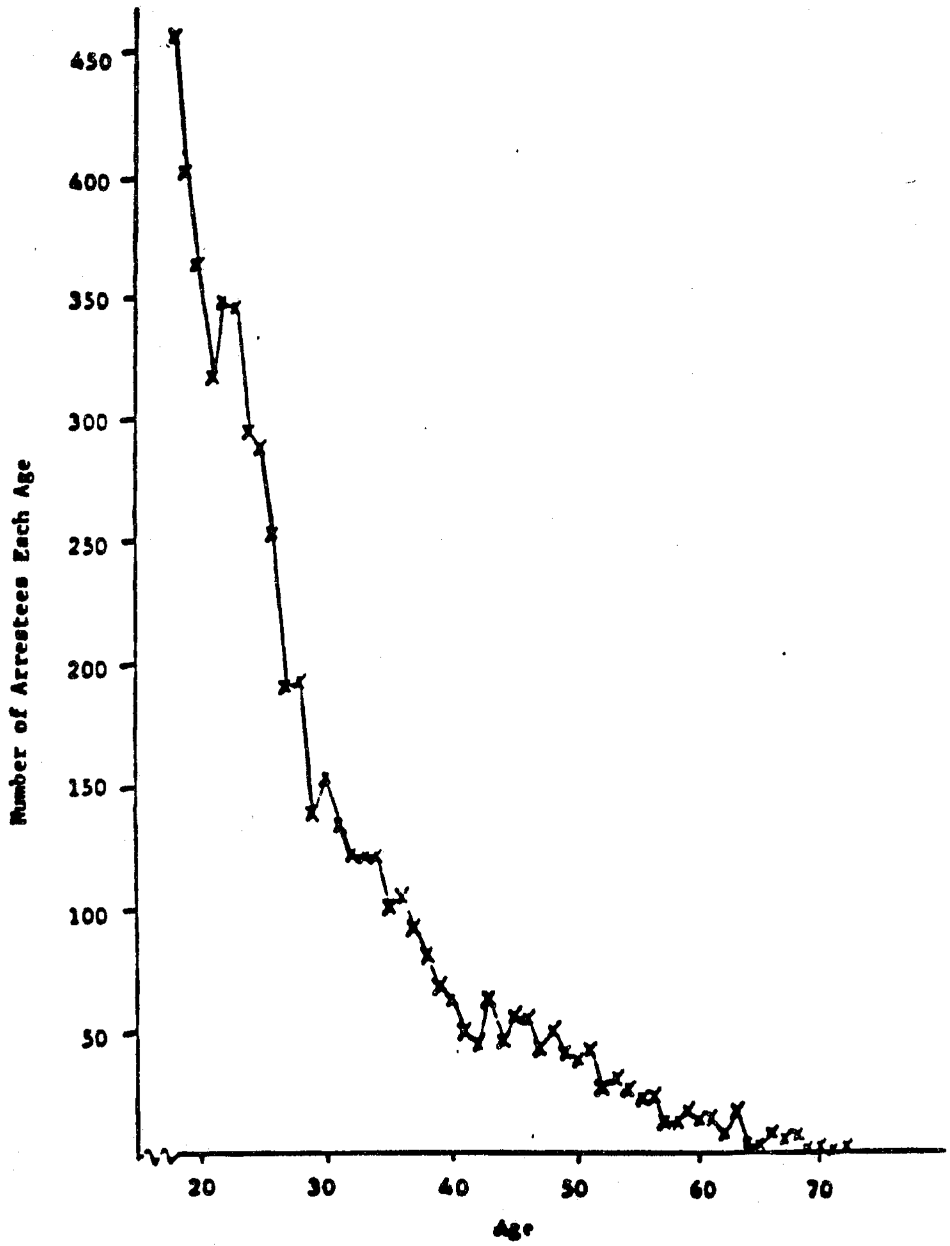


Figure 1: Age Distribution of Criterion Adult Arrestees in Washington, D.C. During 1973

Figure 2 shows that the criminal careers in FBI index offenses for 18 year-old starters can be divided into three periods: 1) the 18-29 year-old "break-in" period characterized by offenders terminating criminal activity at high rates early in careers, 2) the 30-39 year-old "more enduring" period in which few offenders terminate, and 3) the over 40 "burn-out" period during which offenders terminate criminal activity at an increasing rate. In the break-in and burn-out periods, the expected residual adult careers in index offenses were 6 to 7 years, while offenders in their 30's averaged 10 years of remaining criminal activity. In further analyses, Blumstein, Cohen and Hsieh also found that termination rates varied by crime type. Criminal careers involving serious violent crimes (murder, rape, or aggravated assault) are longer and have lower termination rates than careers which include only property offenses (burglary, auto-theft or robbery[8]).

---

[8]Robbery can be viewed as both a violent and a property offense. To the victim robbery is violent since the offender threatens physical harm. However, from the offender's perspective, robbery is also property crime that is committed in order to obtain money. Blumstein, Cohen and Hsieh found that the termination rate for robbery resembled that of property offenses more than violent offenses.

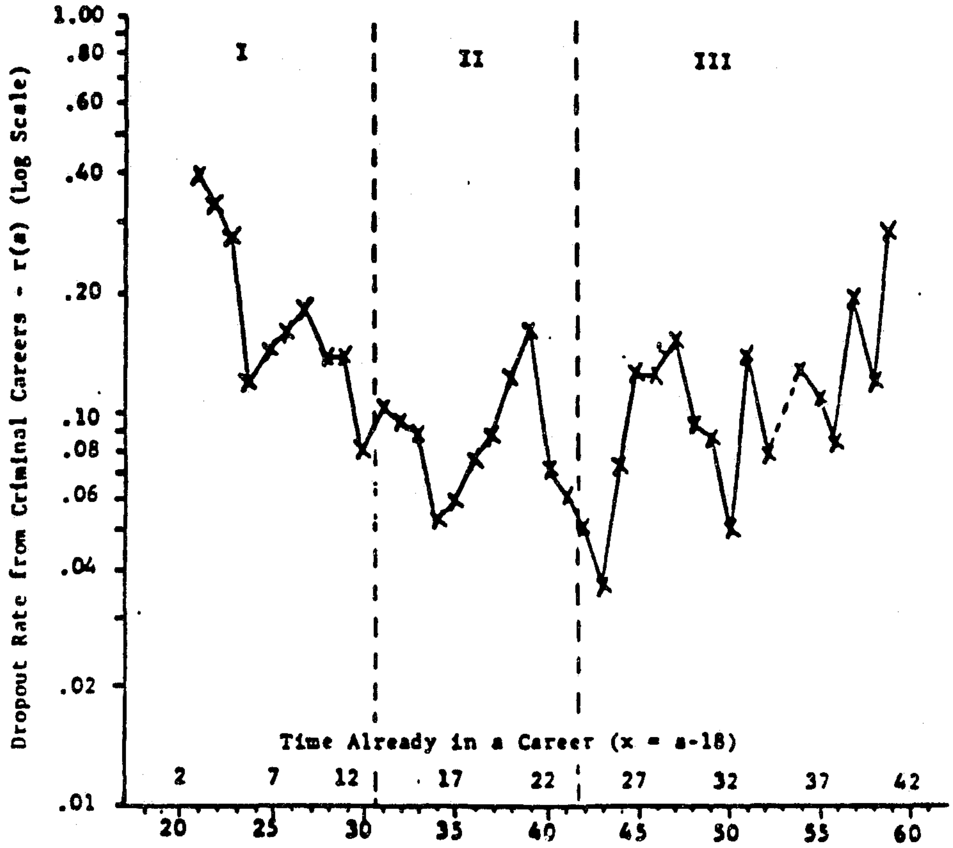


Figure 2: Termination Rate From Criminal Careers by Age for 18-20 Year-Old Starters

In contrast to the life-table approach, several more recent studies have used maximum likelihood techniques to estimate the termination rate. Barnett, Blumstein, and Farrington (1987) modelled the criminal activity of 411 London males from the age of 10 to 24.[9] Their model consisted of two groups of offenders: "Frequents" who had a higher arrest rate μ_1 and a probability P_1 of terminating their criminal career subsequent to a conviction, and "Occasional" with a lower arrest rate μ_2 and a probability P_2 of

[9]A total of 82 out of the 411 youths studied were convicted of one or more criminal offenses.

terminating. They found that the average career length for each group of offenders was between 7 and 9 years. Barnett, Blumstein and Farrington (1988) tested the predictive ability of this model on additional follow-up data for the same offenders between the ages of 25 to 30. Based on 10-24 year-old conviction data, they were able to make accurate predictions of the number of offenders re-convicted, the total number of reconvictions and the average time interval between reconvictions for the offenders from age 25 to 30.

Ahn (1986) modelled the criminal careers of adult arrestees in Detroit and Southern Michigan. The primary focus of his analysis was using hierarchical models to estimate the heterogeneity in the rate at which active offenders are re-arrested. However, in the course of his analysis he also estimated the combined termination rate for all offenders included in his sample under the assumption that career lengths are exponentially distributed. He estimated termination rate as constant at .09 per year, which corresponds to an average career length of 11 years.

1.5 Focus of This Study

The previous analyses of termination rates have generally ignored variations across groups of offenders and focused on estimating overall average termination rates. As was suggested in prior recidivism research, many factors could affect an offender's propensity to terminate criminal offending including, for example: the extent of prior adult and juvenile arrest records on the theory that a more extensive prior record is indicative of greater propensity to continue engaging in crime in the future; employment record since work can be regarded as a diversion from or economic alternative to crime; and history of drug use, since addicts may continue doing crime in order to support their habits.

In the present study, which is based on official arrest histories, only those causal factors that are reflected in an offender's prior adult arrest record can be explored. Using official adult arrest records from the Detroit SMSA[10], each individual's record is divided into an initial period used to identify attributes of the offender's prior record and a follow-up period. Offenders with similar prior records are grouped together and maximum likelihood techniques are used to estimate the average termination rate (δ) for each group from the follow-up records. Then observed variation in δ across groups is then explored in an

---

[10]The data for both this study and Ahn (1986) were drawn from computerized criminal history files maintained by the Federal Bureau of Investigation. The data in the present study, however, is augmented by a more extensive 5 to 8 year follow-up of the sampled arrestees.

effort to identify which combinations of attributes are associated with higher or lower termination rates.

This work expands upon the maximum likelihood analyses of Ahn (1986) and Barnett, Blumstein and Farrington (1988) by focusing on the variation in termination rates across offender attributes. Blumstein, Cohen, and Hsieh (1982) provided the only previous estimates that considered variations in termination rates for offenders. Relying on life-table techniques, however, those estimates required strong assumptions about the expected progress of future offending based on past arrest histories. The present analysis avoids these assumptions by relying on observed arrests during a follow-up period in order to establish δ.

2.0 THE DATA

The data used in this study consists of the adult[11] arrest records of 19,852 males arrested for at least one of the six most serious of the FBI index offenses (designated in this study as "criterion[12]" offenses) in the Detroit SMSA between January 1974 and December 1977. An individual's arrest history includes for each adult arrest the date of arrest, a list of the offenses charged at the time of each arrest, and the final disposition of the arrest (conviction or not). The record includes arrests from age 17 through June, 1982.

2.1 Arrestee Attributes

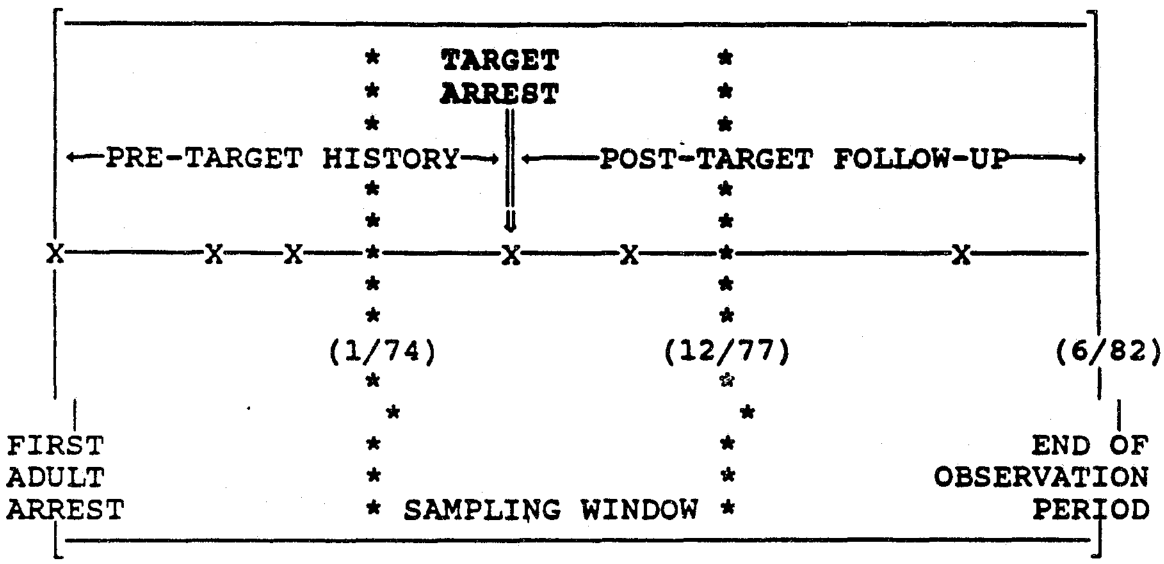
This study focuses on estimating the termination rate following an arrest for groups of offenders with similar attributes. Offender attributes are determined from demographic data and from each offender's criminal record prior to a "target" arrest. An offender's target arrest is defined as his first criterion arrest (arrest for a criterion offense) within the January 1974 to December 1977 window period. Figure 3 displays the division of an offender's criminal record into a pre-target history -- used to establish an offender's attributes -- and a post-target,

---

[11]The age of adult jurisdiction in Michigan is 17.

[12]The criterion offenses include homicide, forcible rape, aggravated assault, robbery, burglary and auto theft. The FBI index offenses of larceny and arson (which was not added to the index offenses until 1979) are excluded from the criterion offenses.

follow-up -- used to estimate the termination rate for offenders with similar attributes.[13]



X indicates an arrest

Figure 3: Pre-Target, Post-Target Partition of an Arrest Record

Each offender's Pre-target arrest history and demographic data are summarized by the eight offender attributes described in Table 1. To compactly describe an offender's attribute values an attribute ID composed of Race, Age, Priors, and Crime Types sections is used. The Race section includes the code value for the arrestee's RACE, Age includes values for AGE1 and AGENOW, Priors includes values for CPRIOR and IPRIOR, and Crime Types includes values for VEVER, REVER and DEVER. For example, a black, 24 year-

[13]Several adjustments to the data are explained in Appendix B.

old, arrestee with 4 prior criterion arrests, two prior incarcerations for criterion offenses, and who is charged with robbery on the target arrest has an attribute ID of B-12-3I-0R0.

Table 1: List of Arrestee Attributes

Attribute	Values	Description
RACE	W. White B. Black	Arrestee's race. Only white and black arrestees are included in the database.
AGE1	1. 17-19 2. 20-29 3. 30+	Arrestee's age at his first adult criterion arrest ever.
AGENOW	1. 17-19 2. 20-29 3. 30-39 4. 40+	Arrestee's age at his target arrest (first criterion arrest between 1/74 and 12/77).
CPRIOR	0. No Arrests 1. 1 or 2(few) 3. 3+ (many)	The number of adult criterion arrests which precede the arrestee's target arrest.
IPRIOR	0. No Prior Incar. I. 1+ Incar.	Indicator of whether the arrestee was ever incarcerated for a criterion arrest prior to his target arrest.
VEVER	0. No Violent V. Violent	Indicator of whether the arrestee was ever charged with a violent offense either prior to or on his target arrest.
REVER	0. No Robbery R. Robbery	Indicator of whether the arrestee was ever charged with robbery either prior to or on his target arrest.
DEVER	0. No Drugs D. Drugs	Indicator of whether the arrestee was ever charged with a drug offense either prior to or on his target arrest.

Each of the prior record attributes is chosen to test the extent to which the data supports theoretically hypothesized influences on criminal career termination. While it is not a relevant theoretical construct, race is included in the set of attributes to determine if offender race is associated with the termination rate and to test whether the influence of other theoretical constructs varies across the races. Previous studies have established that a higher proportion of urban black males than urban white males participate in criminal activity. However, of those black males who do become criminally active the frequency of offending is much more similar between black and white offenders.[14] The question regarding the termination rates is whether those of black and white offenders are similarly close.

Two variable associated with age (AGE1 and AGENOW) are included. Offenders currently in their thirties (AGENOW = 3) are expected to have the lowest termination rate base upon the results of Blumstein, Cohen, and Hsieh (1982). For any particular value of AGENOW, the termination rate is expected to be lower for younger starters on the hypothesis that offenders that have been active for a longer time will be more committed offenders and so are less likely to terminate offending. AGENOW is interpreted as reflecting

---

[14]Blumstein, Cohen, Roth, and Visher (1986) in an analysis of prior literature state that the participation rate among black offenders is higher than among white offenders (pp. 252-253), but that the rate of offending among active offenders does not vary substantially with race (p. 352).

the impact on δ as a function of offenders age at target arrest and not a continuous aging process of δ. Under this assumption, two offenders currently aged 22 need not have the same δ if their age at the target arrests were 18 and 21 (i.e., the offenders differ on AGENOW).

Similarly, the termination rate is expected to decrease as the extent of an offenders' prior criminal record increases as measured by CPRIOR and IPRIOR. This decrease could be associated with either a decrease in individuals' propensity or ability to terminate offending after establishing a criminal record or due to heterogeneity within groups. For example, suppose a group contains two sub-populations with distinct termination rates $\delta_2 > \delta_1$. Then a higher proportion of sub-group 2 offenders would terminate prior to obtaining an extensive arrest record. Therefore, the group of offenders with many prior arrests would have a higher proportion of offenders from sub-group 1 than the group of offenders with no prior arrests, and the average termination rate would exhibit a decrease as the number of prior arrests increased. Along the same lines, those offenders who are still criminally active in spite of a prior incarceration are expected to have a lower termination rate.

Blumstein, Cohen, and Hsieh (1982) observed that careers which included violence on average were longer than those which did not. So, it is expected that δ will be lower if VEVER is positive. Robbery is included separately because of its importance as a serious property crime, but it is often classified as a violent crime (for example, in the Uniform Crime Reports). Blumstein, Cohen, and Hsieh (1982) observed that careers which included robbery were more similar to careers involving only property crimes than careers which involved violent crimes. Therefore, the termination rate is expected to be the same if REVER is positive or not.

DEVER is included as an indicator of a relationship with the drug community and possibly of drug use. Under a hypothesis that a large fraction of people arrested for drugs are heavy users, then this addiction would be expected to be associated with a lower termination rate.

2.2 Offender Groups Available in the Data

An offender group is defined as a collection of offenders with identical attribute ID's. These groups are established in order to estimate the termination rate, δ, for similar offenders. A total of 1,152 combinations of the levels of the 8 offender attributes are possible. However, not all 1,152 combinations represent logically possible offender groups. For example, an arrestee whose first arrest as an adult occurred while he was in his thirties can not have his target arrest while he is in his late teens. Thus, the 1,152 possible combinations of attributes reduce to 704 logically possible groups.[15]

Since the 19,852 arrestees in the database are not uniformly

---

[15]The calculation of the number of logically possible groups is presented in Appendix A.

distributed across groups, only some of the 704 logically possible groups have enough cases to be analyzed. Of the 704 logically possible groups, 478 have one or more arrestees, but only twenty groups have a sample size larger than 150 offenders[16]. Two of these groups -- black and white, 17-19 year-old, first-time arrestees (coded as B-11-00-000 and W-11-00-000) -- account for 20% of all arrestees.

2.3 Aggregation of Groups

In order to increase the number of offender groups in the analysis, smaller groups are combined on the basis of similar attributes. The goal of this aggregation is to combine smaller groups into aggregate groups in order to increase within group sample size, yet maintain similarity of arrestee attributes within the aggregate groups. In particular, aggregate groups are not formed by combining groups which differ across race and rarely across number of prior arrests (CPRIOR). In this manner, 19 groups and 25 aggregate groups including 15,703 arrestees (79% of the 19852 arrestees in the adjusted database) are included in the analysis.

Throughout the remainder of this report, groups and aggregate groups will both be referred to simply as "groups." Groups are

[16]Simulation results presented in Appendix D report that reasonably accurate estimates of δ can be obtained with a sample size as small as 150 offenders.

identified by their attribute ID[17] or, since it is often convenient to use a single number, by their "rank number." The rank number indicates the approximate rank of a group in terms of the number of cases it contains. The group with the largest sample size (n=2700) has rank number 1 and the group with the smallest sample size (n=80) has rank number 44.[18]

Figure 4 displays the groups and aggregate groups in a grid format.[19] The grid columns are defined by demographic attributes, first by RACE (black or white) and within each race by the two digit age identifier (AGE1, AGENOW). The rows of the grid are divided by criminal record attributes, first by the two digit prior record identifier (CPRIOR, IPRIOR), and within each prior value by the 3 digit crime types identifier (VEVER, REVER, & DEVER).

---

[17]The attribute ID used to refer to an aggregate group includes the values of attributes that are shared by all the groups included in the aggregate and a "*" for those attributes whose value varies across the groups in the aggregate.

For example, aggregate group 20 (refer to Figure 4) combines 5 groups of black arrestees with few prior arrests and a charge of violence. These groups vary across age at first arrest, age at target arrest, and whether they were charged with robbery or not. Therefore, this aggregate group's attribute ID is B-**-10-V*0.

[18]The rank number is the relative sample size before cases were deleted from the data as described in Appendix B.

[19]Groups combined to form an aggregate group are identified in Figure G with the same rank.

Figure 4: Arrestee Groups Included in the Analysis

PRIOR & CRIME TYPE CODES		RACE & AGE CODES: WHITE 11	12	13	14	22	23	24	33	34	BLACK 11	12	13	14	22	23	24	33	34
	00 000	1				6			10	14	2				7			34	
	00D	12				15					29				29				
FIRST	0R0	8				25					5				13				
ARREST	0RD	28				28					33				33				
	V00	3				4			10	14	9				11			34	38
	V0D	36				23													
	VR0																		
	VRD																		
	10 000	17	17			17					16	16			16				
	00D	32	32			32									33				
	0R0	44	44								24	24			24				
	0RD		39												33				
FEW	V00	19	19	31		19	31				20	20			20				
PRIOR	V0D		39			39													
ARRESTS	VR0											20			20				
	VRD																		
	1I 000	43	43								26	26			26				
	00D																		
	0R0										18	18			18				
	0RD																		
	V00					43						22			22				
	V0D																		
	VR0											22			22				
	VRD																		
	30 000																		
	00D																		
	0R0																		
	0RD																		
MANY	V00																		
PRIOR	V0D																		
ARRESTS	VR0																		
	VRD																		
	3I 000		30									27			41				
	00D											40							
	0R0		30									35			41				
	0RD											40							
	V00		37	31								21	42						
	V0D																		
	VR0		37	31								21	42		41	42			
	VRD											40							

The following four types of groups are not well represented in the database (see Figure 4).

1. Offenders with long careers are rare: Only 2 groups of arrestees (groups 31 and 42) have a first arrest between ages 17-29 (AGE1 = 1 or 2) and a target arrest after age 29 (AGENOW = 3 or 4). In contrast, 23 of the 44 groups in the study are groups of first-time arrestees (Priors = 00).

2. Black arrestees with a few prior arrests and drug charges are rarer than similar white arrestees: No groups are composed solely of black arrestees with few prior arrests and drug charges. Group 33 contains some Blacks with few prior arrests and drug charges, but it also contains black, first-time arrestees with drug charges. In contrast, two groups of white arrestees with few prior arrests and drug charges are included in the study (groups 32 and 19).

3. White arrestees with few prior arrests who were previously incarcerated are rarer than similar black arrestees: Group 43 is the only group of white arrestees with few prior arrests and a prior incarceration. In contrast, there are three groups of black arrestees with few prior arrests and a prior incarceration (26, 18, and 22).

4. Arrestees with many prior arrests and no prior incarcerations (Priors = 30) are not represented in the sample: It is thus rare for offenders to accumulate many arrests for criterions offenses without being incarcerated.

3.0 THE ANALYSIS

The average termination rate (δ) is calculated for each group of offenders using maximum likelihood estimation. The observed variation in $\hat{\delta}$ across groups is then summarized by a model developed using weighted least squares regression.

3.1 The Model

The termination rate of criminal activity (δ) is estimated for each group of offenders based upon the offenders' post-target follow-up periods. Specifically, each offender has a "next-arrest interval" defined as the time from the offender's target arrest until the offender's next criterion arrest or through June, 1982, (the end of the study) whichever comes first.

In general, a post target follow-up period without a next arrest suggests that the offender may have terminated criminal activity. Absence of a next arrest, however, might also result when an offender remains active, but with a very low arrest rate. In this case, re-arrest may still occur, but not before the end of the observation period. Because of the inherent difficulty of isolating career termination (δ) from low arrest rates (μ) for still active offenders, δ and μ are estimated simultaneously for each group.

In order to estimate δ and μ, the following assumptions are made:

1. Each offender has a time-invariant arrest rate (μ) for the duration of the next-arrest interval. The probability that an active, non-incarcerated offender is arrested in a small time interval (dt) is μ x dt. It is further assumed that offenders are not re-arrested while imprisoned.[20]

2. Each offender has a time-invariant termination rate (δ) for the duration of the follow-up interval. The probability an active offender terminates offending in a small time interval (dt) is δ x dt. It is assumed that offenders can terminate criminal activity even while in prison.[21]

3. The arrest rate (μ) is homogeneous across offenders in a group.

4. The termination rate (δ) is homogeneous across offenders in a group.[22]

---

[20] A constant arrest rate might arise if offenders commit crimes at a fairly constant rate, and if the probability of arrest for each crime is also constant over time. For example, on average a high-rate offender might commit robbery about 5 times a week, and the probability the offender is arrested for any particular robbery might be 1 in a 1,000. On this basis, the offender's arrest rate is (5 $week^{-1}$) x (52 weeks per year) x (1/1,000 arrests per crime) = .26 arrests per year.

In this Detroit sample it has previously been estimated that between 40 and 50% of all arrests are recorded. Despite the substantial level of non-recording, there appears to be no systematic bias that attributes missing arrest records to particular crime types or subgroups of offenders. Thus, the recorded arrests could still provide a representative sampling of all arrests. The impact of non-recording is that estimates of μ are lower in magnitude, while estimates of δ are generally unaffected.

[21] The assumption that μ and δ are time-invariant over an offender's next-arrest interval (at most 8.5 years) is not equivalent to the more restrictive assumption that μ and δ are constant over the offender's entire life-time. Changes in μ and δ as offenders age or accumulate additional arrests are observable in the present analysis through comparisons of the estimates of μ and δ between offender groups with less extensive prior records and those with more.

[22] The appropriateness of the assumption of homogeneous μ and δ within a group is dependent on the extent to which the offender attributes selected for analysis in fact do explain differences in individual offending behavior. Homogeneous μ and δ within a group obviates the need to control for sample selection bias. In the

Based on these assumptions, the probability (or likelihood) of observing the offender not in prison for a period of time F during the next-arrest interval is a function of μ and δ as follows:[23]

$$P(F|J,\mu,\delta) =$$

$$\begin{cases} \Pr\begin{bmatrix}\text{DOES NOT} \\ \text{DROPOUT WHILE} \\ \text{IN JAIL}\end{bmatrix} \times \Pr\begin{bmatrix}\text{RE-ARREST} \\ \text{F MONTHS} \\ \text{LATER}\end{bmatrix}\begin{matrix}\text{DOES NOT} \\ \text{DROPOUT} \\ \text{IN JAIL}\end{bmatrix} & \text{, re-arrest observed} \\ 1 - \Pr\begin{bmatrix}\text{DOES NOT} \\ \text{DROPOUT WHILE} \\ \text{IN JAIL}\end{bmatrix} \times \Pr\begin{bmatrix}\text{RE-ARREST} \\ \text{WITHIN F} \\ \text{MONTHS}\end{bmatrix}\begin{matrix}\text{DOES NOT} \\ \text{DROPOUT} \\ \text{IN JAIL}\end{bmatrix} & \text{, no re-arrest observed} \end{cases}$$

$$= \begin{cases} \mu\ e^{-\delta(J+F)} \times e^{-\mu(F)}\ dt & \text{, re-arrest observed} \\ 1 - \left[\frac{\mu}{\mu+\delta}\ e^{-\delta J} - \frac{\mu}{\mu+\delta}\ e^{-\delta(J+F)}\ e^{-\mu(F)}\right] & \text{, no re-arrest observed} \end{cases}$$

F = number of months free during the follow-up interval

---

homogeneous case, the parameters for offenders who have a target arrest and are thereby included in the sample are the same as for offenders with identical attributes in the general population of offenders from which the sample was drawn. To the extent that the offender attributes do not fully characterize the differences in offending behavior, and substantial heterogeneity in μ and δ remains within the groups studied, the estimates of μ and δ calculated for the sampled population would differ from those of the general population. The population sampled would tend to include more active (higher μ) and more persistent (lower δ) offenders.

[23] See appendix C for the derivation of this formula.

J = number of months spent incarcerated during the incarceration interval.

Combining the likelihoods of next-arrest intervals observed in a group, maximum likelihood estimates of μ and δ are obtained for each group.

3.2 Maximum Likelihood Estimates

Figures 3-1 and 3-2 present the maximum likelihood estimates of δ and μ, respectively, for each group.[24] Figure 3-1 shows a considerable range of variation in δ across offender groups. The highest termination rates -- which correspond to the shortest expected residual careers -- are found among white, first-time arrestees. A typical value of δ among these groups is .19 per year (or an expected residual career of $1/\delta \approx 5$ years). Many of the lowest termination rates, typically $\delta \approx .05$ per year or $1/\delta \approx 20$ years, are observed among both white and black arrestees who have extensive prior records.

The patterns of variation observed in μ across the offender groups are less easily distinguished. Consistent with previous results, μ is similar in magnitude across race and over prior record, and is lower for offenders who are active in violent offenses (Blumstein and Cohen, 1979; Blumstein et. al., 1986).

[24]These two figures use the same format as the grid in Figure 4. Appendix E reports these same results in tabular form.

Figure 3-1: Variation in Annual Dropout Rates, δ, by Offender Group

PRIOR & CRIME TYPE CODES			RACE & AGE CODES: WHITE 11	WHITE 12	WHITE 13	WHITE 14	WHITE 22	WHITE 23	WHITE 24	WHITE 33	WHITE 34	BLACK 11	BLACK 12	BLACK 13	BLACK 14	BLACK 22	BLACK 23	BLACK 24	BLACK 33	BLACK 34
	00	000	.19				.23			.05	.21	.05				.09			.11	
		OOD	.14				.22					.10				.10				
FIRST		ORO	.14				.23					.09				.05				
ARREST		ORD	.22				.22					.07				.07				
		VOO	.17				.11			.05	.21	.07				.10			.11	.55
		VOD	.18				.04													
		VRO																		
		VRD																		
	10	000	.12	.12			.12					.08	.08			.08				
		OOD	.12	.12			.12									.07				
		ORO	.15	.15								.08	.08			.08				
		ORD		.11												.07				
FEW		VOO	.10	.10	.05		.10	.05				.07	.07			.07				
PRIOR		VOD		.11			.11													
ARRESTS		VRO											.07			.07				
		VRD																		
	1I	000	.11	.11								.10	.10			.10				
		OOD																		
		ORO										.05	.05			.05				
		ORD																		
		VOO					.11						.04			.04				
		VOD																		
		VRO											.04			.04				
		VRD																		
	30	000																		
		OOD																		
		ORO																		
		ORD																		
MANY		VOO																		
PRIOR		VOD																		
ARRESTS		VRO																		
		VRD																		
	3I	000		.03									.09			.11				
		OOD											.14							
		ORO		.03									.01			.11				
		ORD											.14							
		VOO		.08	.05								.02	.04						
		VOD																		
		VRO		.08	.05								.02	.04		.11	.04			
		VRD											.14							

Figure 3-2: Variation in Annual Arrest Frequencies, μ, by Offender Group

PRIOR & CRIME TYPE CODES		RACE & AGE CODES																	
		WHITE									BLACK								
		11	12	13	14	22	23	24	33	34	11	12	13	14	22	23	24	33	34
00	000	.25				.14			.04	.03	.31				.20			.07	
	00D	.23				.24					.24				.24				
FIRST	0R0	.18				.12					.26				.16				
ARREST	0RD	.19				.19					.26				.26				
	V00	.18				.08			.04	.03	.23				.12			.07	.05
	V0D	.19				.08													
	VR0																		
	VRD																		
10	000	.26	.26			.26					.34	.34			.34				
	00D	.28	.28			.28									.26				
	0R0	.24	.24								.28	.28			.28				
	0RD		.25												.26				
FEW	V00	.11	.11	.08		.11	.08				.18	.18			.18				
PRIOR	V0D		.25			.25													
ARRESTS	VR0											.18			.18				
	VRD																		
1I	000	.27	.27								.37	.37			.37				
	00D																		
	0R0										.26	.26			.26				
	0RD																		
	V00					.27						.23			.23				
	V0D																		
	VR0											.23			.23				
	VRD																		
30	000																		
	00D																		
	0R0																		
	0RD																		
FEW	V00																		
PRIOR	V0D																		
ARRESTS	VR0																		
	VRD																		
3I	000		.23									.40			.34				
	00D											.38							
	0R0		.23									.18			.34				
	0RD											.38							
	V00		.22	.08								.20	.11						
	V0D																		
	VR0		.22	.08								.20	.11		.34	.11			
	VRD											.38							

3.3 Clusters of Groups

The termination rate estimates show considerable variation across groups with rates ranging from .02 to .55 per year. Figure 3-1 indicates that there is regularity in this variation such that groups with similar attributes often have similar termination rates. For example, each of the groups of 17-19 year-old, first-time, white arrestees exhibited a termination rate between .14 and .22 per year, irrespective of the type of crimes in which they engaged.

Weighted least-squares regression is used to determine the attributes for which significant variation in $\hat{\delta}$ is observed and, thereby, to identify clusters of groups. A cluster is defined as a collection of offender groups with both similar (although not identical) attributes and similar termination rates.

The following regression model summarizes the observed variation in $\hat{\delta}$ across groups(the standard errors of estimated coefficients are provided in parentheses):[25]

(3-1)

$$\hat{\delta} = \underset{(.01)}{.18} - \underset{(.01)}{.11}\, X_b + \underset{(.11)}{.09}\, X_1 - \underset{(.02)}{.06}\, X_{wf} - \underset{(.03)}{.13}\, X_{wm} - \underset{(.03)}{.10}\, X_{wtv} + \epsilon$$

$R^2 = .7650$

[25]The regression model was chosen for both amount of variation in $\hat{\delta}$ explained and conciseness. A discussion of the procedure used to select this model is provided in Appendix G. All coefficients are significantly different from zero at the α=.01 level, except one. The parameter X_l is included in the model due to its large magnitude in spite of its lack of statistical significance.

Where,

$$X_b = \begin{cases} 1 & \text{black and age} < 40 \\ 0 & \text{else} \end{cases}$$

$$X_1 = \begin{cases} 1 & \text{aged 40+} \\ 0 & \text{else} \end{cases}$$

$$X_{wf} = \begin{cases} 1 & \text{white, few prior arrests, age} < 40 \\ 0 & \text{else} \end{cases}$$

$$X_{wm} = \begin{cases} 1 & \text{white, many prior arrests, age} < 40 \\ 0 & \text{else} \end{cases}$$

$$X_{wtv} = \begin{cases} 1 & \text{white, no prior arrests, violent, aged 20-39} \\ 0 & \text{else} \end{cases}$$

This model shows that 5 binary parameters can account for most (76%) of the observed variation in $\hat{\delta}$ across groups. Furthermore, this model identifies those attributes associated with observed variation in $\hat{\delta}$: offenders aged 40+ exhibited a higher termination rate than those aged less than 40; among 17-39 year-old white offender groups the termination rate exhibited a decrease as the number of prior criterion arrests increased; and black 17-39 year-olds, offender groups exhibited a termination rate similar to white offender groups with many prior arrest.

Figure 3-3, a histogram of $\hat{\delta}$ across groups, confirms the conclusion that many white offender groups exhibited a termination rate higher than black offender groups and that the two groups of offenders both white and black aged 40+ exhibit relatively high termination rates. However, the significance of this finding with respect to offenders aged 40+ is limited by the small sample (only 2 groups) of offenders of this type.

****** Figure 3-3 here*****

Among 17-39 year-old white offenders, $\hat{\delta}$ is observed to decrease significantly as the number of prior arrests increases. Figure 3-4 reports that first-time white arrestees (denoted by O) exhibit a relatively high termination rate, a lower rate for offenders with few prior arrests (denoted by F), and the lowest rate for offenders with many prior arrests (denoted by M). White, violent, first-time offenders aged 20-39 years (denoted by O) exhibited a particularly low termination rate, much lower than 17-19 year-old first-time arrestees.

***** Figure 3-4 here *****

Similar differentiation in $\hat{\delta}$ by prior arrests is not observed for black offenders. Figure 3-5 reports that $\hat{\delta}$ for black offender groups is generally in the .07 to .11 per year range and that the variation in $\hat{\delta}$ is not systematically associated with the number of prior arrests.

***** Figure 3-5 here *****

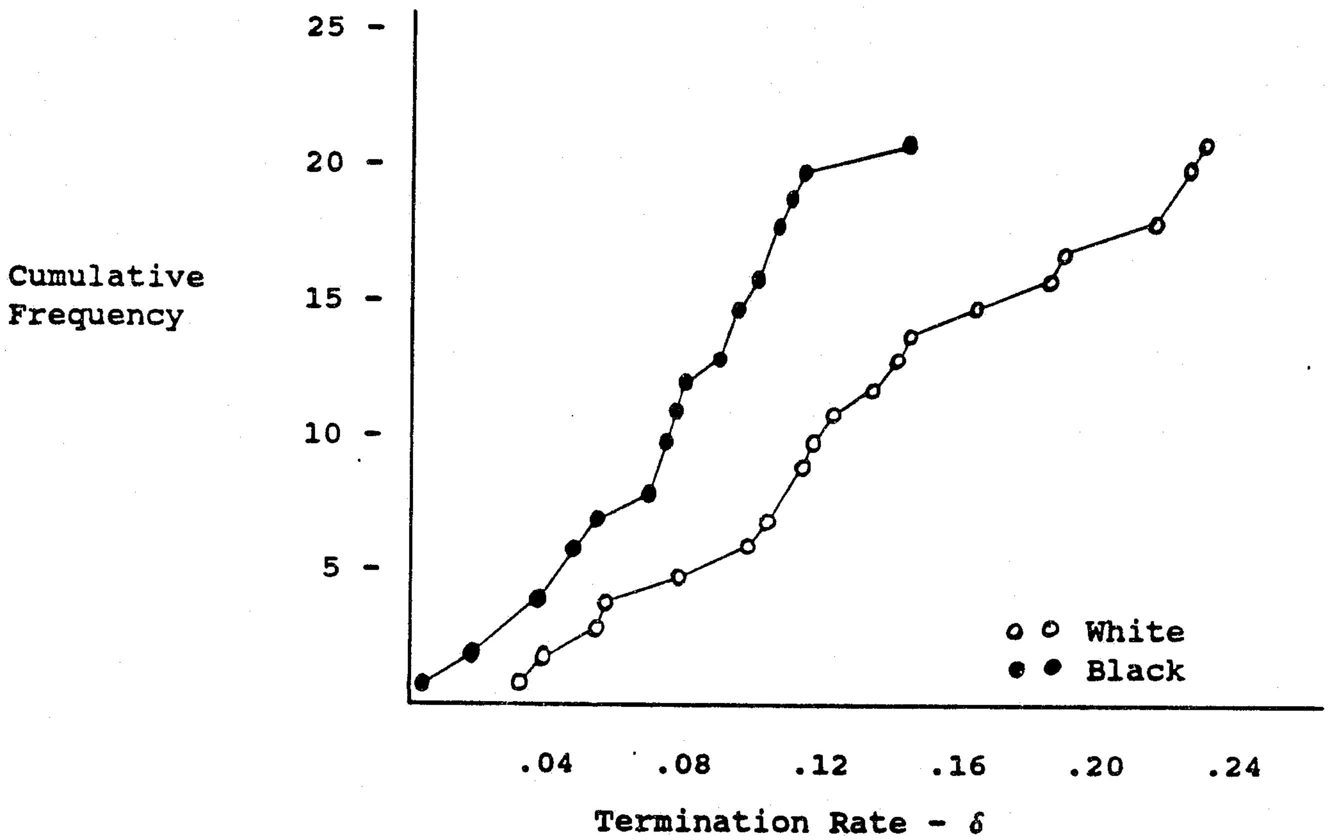


Figure 3-3: Distribution of δ Estimates for White and Black Offender Groups

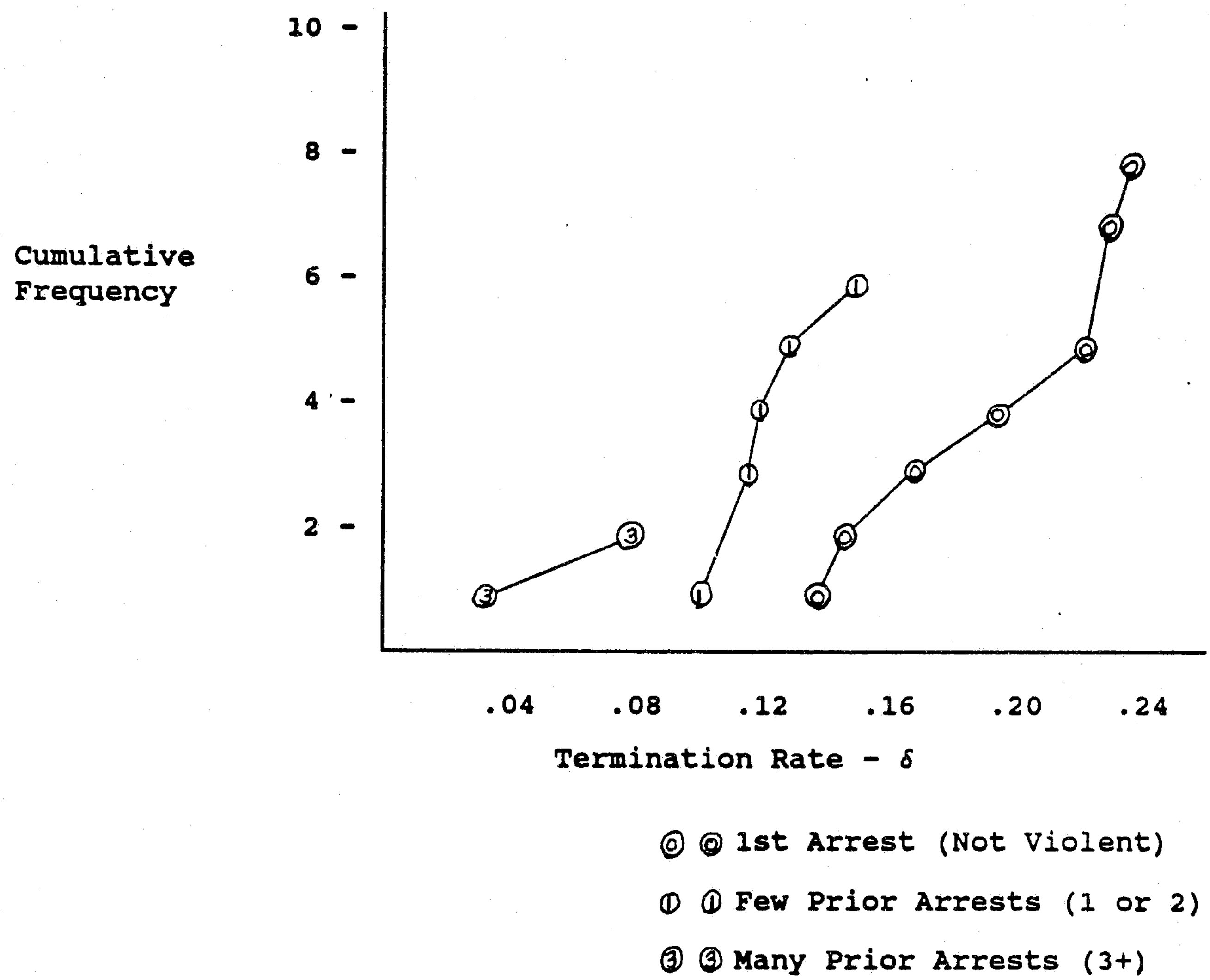


Figure 3-4: Distribution of δ Estimates by Number of Prior Arrests for Criterion Offenses (White Offender Groups Aged 17-39)

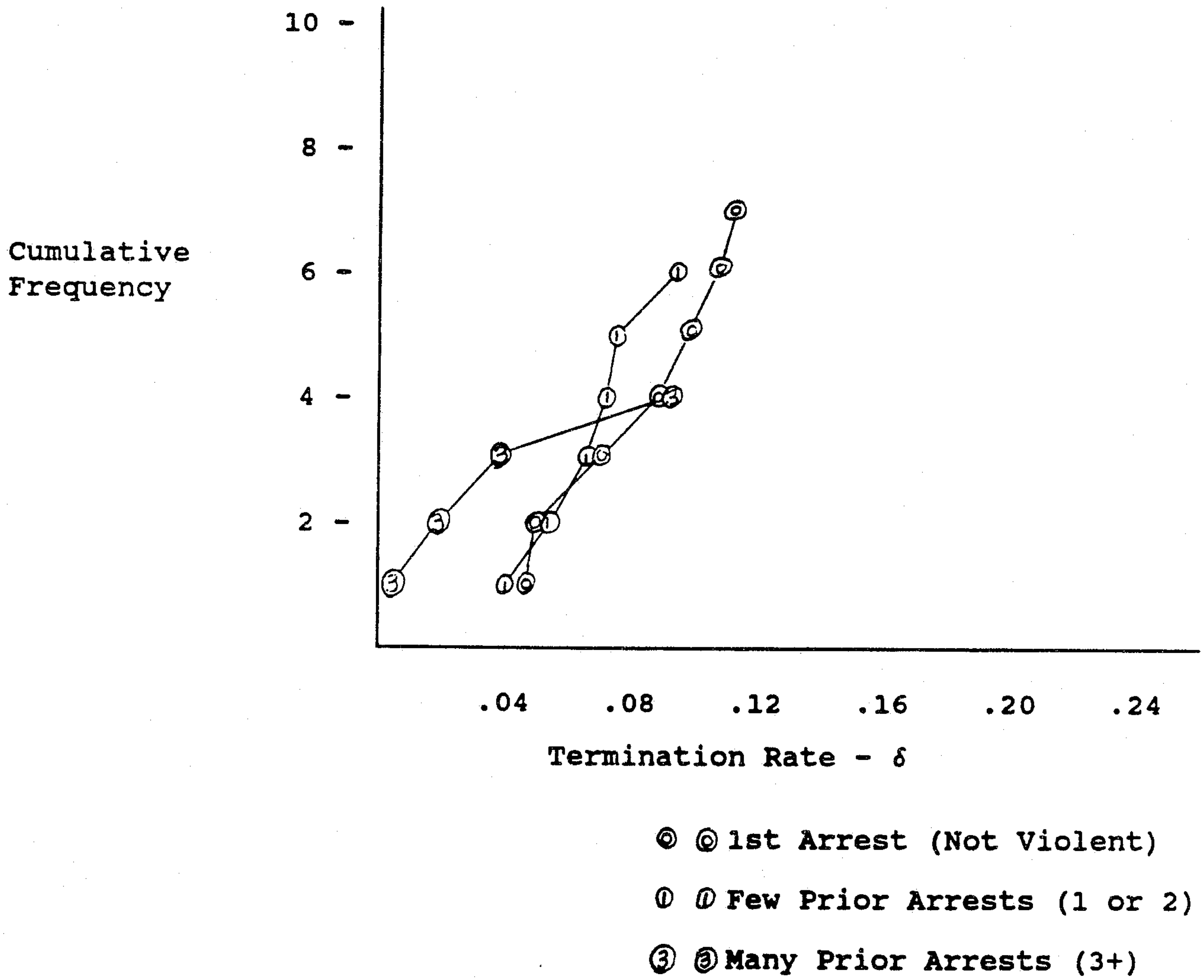


Figure 3-5: Distribution of δ Estimates by Number of Prior Arrests for Criterions Offenses (Black Offender Groups Aged 17-39)

On the basis on this analysis 6 clusters of offender groups with similar attributes and similar values of $\hat{\delta}$ are identified: offenders aged 40+, black offenders aged 17-39, white offenders aged 20-39 with no prior arrests who are arrested for violent offenses, other white 17-39 year-old first-time offenders, white 17-39 year-old offenders with few prior arrests, and white 17-39 year-old offenders with many prior arrests.[26]

[26]See Appendix F for a list of offender groups included in each cluster.

4.0 CONCLUSION

Based upon the regression model for $\hat{\delta}$ (equation 3-1), clusters of offender groups with both similar attributes and similar termination rates can be identified. The analysis of official arrest data for adult male arrestees in the Detroit SMSA yields six clusters of offenders who have distinct attributes and distinct career termination rates.

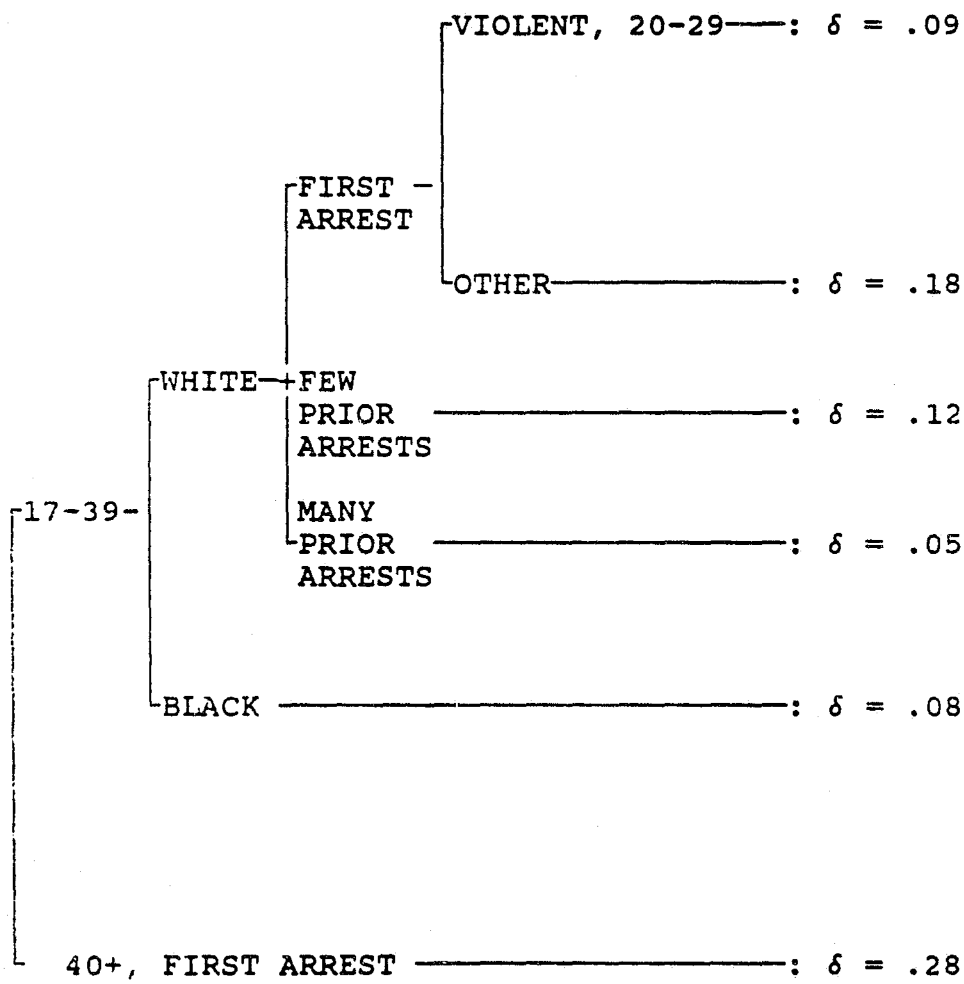


Figure 4-1: Termination Rates for Distinct Offender Clusters

The termination rates for these clusters vary between .05 and .28 (Figure 4.1), and the approximate expected residual career length ($1/\delta$) ranges from 3.5 to 20 years,[27] depending on offender attributes. The variation in δ is related to age, and among 17-29 year-old offenders δ varies with race, and among white arrestees δ varies with the number of prior arrests. On average, offenders with the longest remaining careers (lowest δ) are found in the following clusters:

- 30-39 year-old offenders,
- black, 17-29 year-old offenders,
- white, 17-29 year-old offenders with many prior arrests
- white, 20-29 year-old, first-time, violent offenders.

Two other clusters of offenders are distinguished by their higher termination rates and shorter remaining careers:

- white, 17-29 year-old, first-time arrestees (except 20-29 year-old violent offenders),
- arrestees age 40 or older.

The variation in δ with respect to the eight individual attributes is as follows:

RACE: Black, 17-29 year-old, first-time offenders exhibited a lower termination rate than similar white offenders. No other groups exhibited significant differences in δ across race.

---

[27]An offender's expected residual career length equals $1/\delta$ on the assumption of a constant termination rate for the remainder of the criminal career.

AGE1: A distinction in termination rate based on offender's age at first adult criterion arrest was not observed, but a very limited number of groups differed across this attribute alone.

AGENOW: Offenders in their thirties exhibited a lower termination rate than 17-29 and 40+ year-old offenders.

CPRIOR: Groups which differ across number of prior arrests were available in the dataset primarily among 17-29 year-old offenders. White, 17-29 year-old offenders exhibited a decrease in δ as the number of prior arrests increased. Black, 17-29 year-old offenders did not exhibit a similar pattern.

IPRIOR: Offenders who had previously been incarcerated exhibited termination rates comparable to offenders who had not been incarcerated.

VEVER: Offenders with a prior record of violent offenses generally exhibited termination rates comparable to offenders with no violent prior arrests. The 20-29 year-old, first-time, violent offenders, however, exhibited a lower termination rate than similar not violent offenders.

REVER: Offenders with a prior record of robbery exhibited similar termination rates to non-robbery offenders.

DEVER: Offenders with a prior record of drug-related offenses exhibited termination rates similar to offenders without prior drug-related arrests.

Race-Specific Findings

The termination rate for black, 17-29 year-old offenders does not exhibit the variation with number of prior arrests for criterion offenses (CPRIOR) that is observed for white, 17-29 year-old arrestees (Figures 3.4 and 3.5). White offenders aged 17 to 29 who are arrested for the first time for criterion offenses are distinguished from other offenders the same age by their much shorter remaining careers (higher δ). Once those white offenders with very short careers have dropped out, however, the white and black offenders who continue to remain criminally active are much more similar in their termination rates.

The racial difference in termination rates for first-time arrestees is not likely to be due to differences in police arrest practices. No strong racial differences in arrest risk per crime, especially for the more serious criterion offenses have been found: when the racial mix of arrestees is compared to the racial mix of offenders reported by victims (Hindelang, 1978 and 1981; Messner and South, 1986); when observational data of individual police-citizen encounters are analyzed (Reiss, 1971; Gottfredson and Gottfredson, 1980; Smith, 1984; Gove, et. al., 1985); when city-level data on arrests and crimes by race are used (Liska, et. al., 1985); and when individual-level data on the self-reported crimes and arrests of inmates are used (Petersilia, 1983; Blumstein, et. al., 1988). These consistent results found in a variety of analyses using very different data strongly support the conclusion that observed differences in arrests by race are reasonable proxies of differential involvement in offending by race for young adult, first-time arrestees.

Age-Specific Findings

Offenders in their thirties exhibit a lower termination rate (and, therefore, longer remaining careers) than 17-29 or 40+ year-old offenders. Figure 4.2 shows the termination rates for clusters of offenders as age at target arrest (AGENOW) varies. Those offenders in their thirties who according to conventional wisdom were expected to terminate their offending very shortly, were in fact found to be, on average, the most persistent offenders.

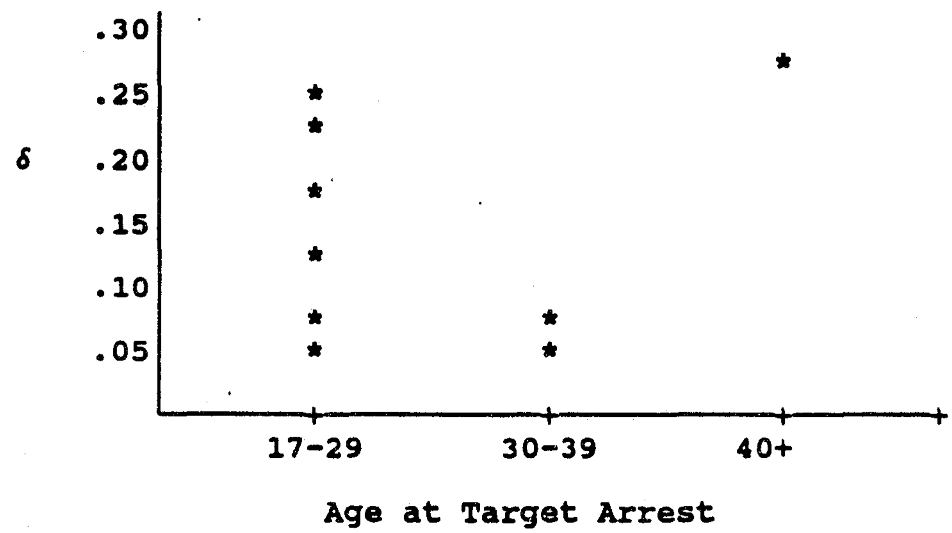


Figure 4.2: Termination Rates of Clusters by Age at Target Arrest

The lower termination rate exhibited by offenders in their thirties is consistent with previous results on residual careers derived using life-table techniques (Blumstein, Cohen and Hsieh, 1982). The fact that maximum likelihood estimates using longitudinal follow-up data for arrestees obtained similar results to previous life-table estimates provides important confirmation of the validity of the life-table approach. This confirmation opens the possibility of much wider use of life-table estimates of termination rates in other jurisdictions since the data needed for life-table estimates -- primarily annual data on the number of arrestees by age -- are much more readily obtained than are individual longitudinal arrest histories like those used in the current analysis.

5.0 AREAS FOR FURTHER RESEARCH

Before the results on termination rates can be used effectively in developing crime control policies further research is required. First, the generality of the variations in termination rates with offender attributes obtained in this study must be tested through replication in other jurisdiction, and perhaps in other time periods. Aside from general replication additional efforts should be made on further refining the list of candidate offender attributes.

The choice of attributes included in this study was largely limited by the information available in the criminal history data. Some additional sources of information that may be related to termination rates include:

- juvenile arrest records (pre-17 year-old data),
- employment and earnings histories,
- information about current drug use (perhaps results of urinalysis at arrest),
- history of marital status.

Data on juvenile criminal records, for example, would permit analysis of the role of early age of onset on termination rates. Data on drug use from urinalysis would permit more accurate identification of those offenders who use drugs than does a history of arrests for drug-related offenses. Employment and earnings history would be an indication of the availability of legitimate opportunities of income as an alternative to continued criminal activity. Marital status would allow for estimation of the potential neutralizing effect on criminal behavior associated with having a spouse.

A refinement of the offender attributes that are considered may help clarify some of the findings in the current analysis. For example, the current analysis failed to find any differences in termination rates for black, 17-29 year-old arrestees. This result for black offenders is intriguing since the termination rates of white offenders are distinguished by several attributes, most notably prior record of arrests. Even this variation in termination rates with prior arrests for white offenders deserves further analysis. Of particular interest is the potential role of the as yet unidentified attributes in accounting for these results. For example, white and black offenders may be sharply distinguished in terms of their juvenile criminal records. If it were the case that juvenile records are far more common among black arrestees, the adult record alone would be a very poor indicator of variations in prior arrests for black adult offenders.

Further analyses are also required to distinguish between the two competing explanations for a decrease in termination rates with increases in the number of prior arrests. First, a decrease in δ with prior arrests might reflect a change in termination rates for individual offenders, whereby an individual's commitment to offending increases with the accumulation of each additional arrest. Alternatively, a decrease in δ with prior arrests might be due to heterogeneity in the mix of offenders at each value of prior record. For example, offenders might be divided along some as yet unidentified attribute into two sub-groups, one with a higher and the other a lower termination rate. This unidentified attribute is related to prior arrests in that as the number of arrests increases, more of the higher termination rate offenders will have ended their criminal activity leaving a lower average termination rate among remaining sample. Thus the decrease in the termination rate with prior arrests could result from a changing mix of offenders or from a change in the termination rates for individual offenders. Distinguishing between these alternatives requires consideration of a wider variety of offender attributes.

Bibliography

Ahn, C., "Hierarchical Stochastic Modelling of Arrest Careers," Ph.D.Thesis, Carnegie-Mellon University, Dept. of Statistics, September, 1986.

Barnett, A., "Prison Populations: A Projection Model," **Operations Research**, Vol. 35, No. 1, January-February, 1987.

Barnett, A., Blumstein, A., and Farrington, D., "Probabilistic Models of Youthful Criminal Careers," **Criminology**, Vol. 25, No. 1, 1987, pp. 83-107.

Barnett, A., Blumstein, A., and Farrington, D., " A Prospective Test of a Criminal Career Model," Unpublished Manuscript, 1988.

Beck, A. and Shipley, B., "Recidivism of Prisoners Released in 1983," **Bureau of Justice Statistics Special Report**, March, 1989.

Blumstein, A., and Cohen, J., "Estimation of Individual Crime rates from Arrest Records," **Journal of Criminal Law & Criminology** Vol. 70, No. 4, 1979, pp. 561-585.

Blumstein, A., and Cohen, J., "Estimating the Duration of Adult Criminal Careers," **Proceedings of the International Statistical**

Institute, Amsterdam, The Netherlands, August, 1985.

Blumstein, A., Cohen, J., and Hsieh, P., "The Duration of Adult Criminal Careers: Final Report to the National Institute of Justice," June, 1982.

Blumstein, A., Cohen, J., Roth, J. and Visher, C., editors, **Criminal Careers and Career Criminals**, National Academy Press, 1986.

Blumstein, A., Cohen, J., and Visher, C., "Linking the Crime and Arrest Process to Measure Individual Crime Rates: Variations in Individual Arrest Risk Per Crime," Unpublished Working Paper, School of Urban and Public Affairs, Carnegie Mellon University, June, 1988.

Gottfredson, M., and Gottfredson, D., **Decision Making in Criminal Justice**, Ballinger Press, 1980.

Gove, W., Hughes, M., and Geerken, M., "Are Uniform Crime Reports a Valid Indicator of The Index Crimes? An Affirmative Answer with Minor Qualifications." **Criminology**, Vol. 23, 1985, pp. 451-501.

Greenwood, P., and Abrahamse, A., **Selective Incapacitation**, Report R-2815-NIJ, Rand Corporation, 1982.

Hindelang, M., "Race and Involvement in Common-Law Personal Crimes." **American Sociological Review**, Vol.43, 1978, pp. 93-109.

Hoffman, P. and Beck, J., "Revalidating the Salient Factor Score: a Research Note," **Journal of Criminal Justice**, August, 1980

Kalbfleisch, J., and Prentice, R., **The Statistical Analysis of Failure Time Data**, John Wiley and Sons, 1980.

Kuester, J., and Mize, J., **Optimization Techniques with Fortran**, McGraw-Hill, 1973.

Lawless, J., **Statistical Models and Methods for Lifetime Data**, John Wiley and Sons, 1982.

Liska, A., Chamlin, M., and Reid, M., "Testing the Economic Production and Conflict Models of Crime Control." **Social Forces** Vol. 64, 1985, pp. 119-138.

Luenberger, D., **Introduction to Linear & Non-Linear Programming**, Addison-Wesley, 1973.

Maltz, M., **Recidivism**, Academic Press, 1984.

Messner, S., and South, S., "Estimating Race-Specific Offending Rates: An Intercity Comparison of Arrest Data and Victim Reports." Unpublished Working Paper. Department of Sociology, State University of New York at Albany, 1986.

Olsson, P., "A Sequential Simplex Program for Solving Minimization Problems," **Journal of Quality Technology**, pp. 53-57, June, 1974.

Petersilia, J., **Racial Disparities in the Criminal Justice System**, Rand Report R-2947-NIC, 1983.

Reiss, A., **The Police and the Public**, Yale University Press, 1971.

Rhodes, W., Tyson, H., Weekley, J., Conley, C., and Powell, G., **Developing Criteria for Identifying Career Criminals**, U.S. Department of Justice, 1982.

Schmidt, P., and Witte, A., "Predicting Criminal Recidivism Using Split Population Survival Time Models," **Journal of Econometrics**, Vol. 40, 1989, pp. 141-159.

Schmidt, P., and Witte, A., **Predicting Criminal Recidivism Using Survival Models**, Springer-Verlag, 1988.

Sechrest, L., White, S., and Brown, E., editors, **The Rehabilitation of Criminal Offenders: Problems and Prospects**, national Academy of Sciences Press, 1979.

Shinnar, R. and Shinnar, S., "The Effect of The Criminal Justice System on the Control of Crime: A Quantitative Approach," **Law and Society Review**, 9:581-612, 1975.

Smith, D., "The Organizational Context of Legal Control." **Criminology**, Vol. 22, 1984, pp. 19-38.

Appendix A: Logically Possible Groups

The logical restrictions on the attributes of a group are due to the chronology of events. Since an offender's first criterion arrest must precede or be at the same age as the target criterion arrest, AGE1 must be less than or equal to AGENOW. If AGE1 is less than AGENOW, then the offender has prior arrests, and CPRIOR (the number of prior arrests) cannot be 0. Finally an offender can only have an incarceration for a prior criterion arrest if he also has a prior criterion arrest. Therefore, if CPRIOR = 0, then IPRIOR = 0. No logical restrictions apply to combining RACE, VEVER, REVER, and DEVER with any other attributes.

The following calculation shows the possible combinations of AGE1, AGENOW, CPRIOR, and IPRIOR. The result in a total of 704 possible combinations of attributes that identify logically possible groups of offenders.

AGE1	AGENOW	(CPRIOR, IPRIOR Combinations)	No. of Combinations
1	1	00 10 11 30 31	5
1	2	10 11 30 31	4
1	3	10 11 30 31	4
1	4	10 11 30 31	4
2	2	00 10 11 30 31	5
2	3	10 11 30 31	4
2	4	10 11 30 31	4
3	3	00 10 11 30 31	5
3	4	10 11 30 31	4
4	4	00 10 11 30 31	5
		TOTAL	44

Combinations of Combinations of

<u>(RACE VEVER REVER DEVER)</u>		<u>(AGE1 AGENOW CPRIOR IPRIOR)</u>	
2 x 2 x 2 x 2	x	44	= <u>704</u>

Appendix B: Adjustments to the Data

Several adjustments to the original arrest history data are necessary in order to appropriately estimate μ and δ. The first set of adjustments provides an estimate of when each offender is incarcerated. Since offender's can only be arrested while free (not incarcerated), failure to include incarceration time would unduly bias the estimates of the arrest rate downward.[28] The second set of adjustments remove inappropriately short one-month follow-up intervals until the next criterion arrest or the end of the study (June, 1982), whichever comes first. The inclusion of these inappropriate, one-month, follow-up intervals would unduly bias the estimates of the arrest rate upwards.

Estimates of Incarceration Periods

The original data does not include information about exactly how long each offender spent incarcerated following a conviction. As an approximation, incarceration periods are treated as starting in the month immediately following the month of arrest and continuing for a length of time equal to some fraction of the length of the incarceration sentence imposed on the offender[29] or

---

[28]In this case, the bias to μ is systematic in that a disproportionate number of short inter-arrest intervals are included in the data. The systematic nature of this bias in the estimate of μ also induces a bias in the estimate of δ.

[29]Time served was estimated by the Michigan Department of Corrections as a function of the minimum sentence imposed in court. For sentences of 30 months or less time served average 75% of the minimum sentence length imposed. For sentences over 30 months, the fraction of the sentence actually served decreased as the minimum sentence increased, and time served was estimated as 2.53 + .69 x

until the offender's next arrest[30], whichever comes first.

Offender's whose estimate of time served exceeds June, 1982, (end of the study) and who do not have any follow-up arrests are estimated to be incarcerated for the entire follow-up period. Since there is no time free when these offenders could have committed an offense and been arrested, their records contain no information about how long these offenders were active or whether they terminated their criminal careers. Thus, these offenders, who comprise 1.3% of the original sample, were deleted from the database.

One-Month Follow-up Intervals

Two adjustments are made to the data to control for inappropriate, one-month, follow-up intervals. The first adjustment is the elimination of duplicate recordings of a single arrest. The database is a compilation of arrest information reported from many criminal justice agencies. Duplicate recordings sometimes result when two agencies report the same arrest as having occurred, but in different months, quite often 1 month apart. This disparity can result from uncertainty about the offense date, or a delay between the date when the crime is committed and the date the arrest is made. An arrest is treated as a duplicate report and is dropped from an offender's record, whenever an arrest with the

---

MIN - .00073 x MIN^2.

[30]Time served ends at either a next criterion or non-criterion arrest.

same charges is reported in the offender's record as occurring in the same or preceding month.

The second adjustment is for incarceration terms which extend from the target arrest until the month preceding the next criterion arrest. While these follow-up intervals are likely to be longer than one month, the amount of time spent free during the follow-up period is estimated to be only one-month long. Many of these one-month intervals reflect imprecision in the time served estimate that results in time served overlapping the time to the next arrest. When this occurs, the incarceration period is adjusted downward to end in the month before the next arrest yielding one-month intervals.

Some of these arrests before the end of the time served interval may be due to arrest double-counts in which the difference in arrest dates is larger than one month. Other arrests may be due to arrests for crimes the individual committed before the target arrest that were not charged to the offender until after being incarcerated. In either of these cases, a next arrest after only one month free represents an anomaly of the data collection procedures and not evidence of the timing of actual follow-up offending. Offender records in which follow-up intervals of only one month free follow time spent incarcerated, which comprise 4.5% of the original sample, are treated as potential data collection anomalies and are deleted from the database.

Appendix C: Derivations

This appendix is composed of four sections of derivations. In the first section the basic probabilistic analysis associated with μ and δ is introduced. In this manner, many of the probabilistic quantities used in the likelihood function are derived. The second section is composed of the derivation of the likelihood function for an offender's follow-up interval. In the third section a the assumptions of the model are validated using a χ^2 test. The fourth section is composed of the derivation of a maximum likelihood estimator for a homogeneous population under idealized conditions. The fifth section analyses the result of using this idealized one-population maximum likelihood estimator when the data is from a two-population sample. The results of the fifth section are used in evaluating the results of simulation analyses (Appendix D).

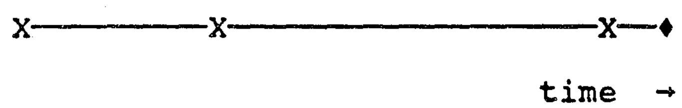
Probabilistic View of Criminal Careers

An offender's criminal record can be viewed as the result of two processes which occur in parallel: the arrest process and the termination process. Under the assumption of a time-invariant termination rate, the distribution of career length is exponentially distributed with an expected career length of $1/\delta$.[31] Under the assumption of a time-invariant arrest rate, the distribution of inter-arrest intervals without consideration of career termination is exponentially distributed with expected

[31] For a more complete discussion of the calculus of duration data see Lawless, pp. 8-10.

inter-arrest time of $1/\mu$.

The arrest and termination processes can be viewed as a combined process of competing events where an event is defined as either an arrest or career termination.[32] This process results in a series of arrests which ends with a career termination event as illustrated in Figure C-1.



X indicates an arrest

♦ indicates career termination

Figure C-1: Example of a Criminal Career

This combined process is characterized by the distribution of inter-event times and the probability an event is an arrest. The inter-event times for a combination of processes with exponentially distributed inter-event times and expected inter-event times of $1/\mu$ and $1/\delta$ is also exponentially distributed and has an expected inter-event time of $1/(\mu+\delta)$.[33] The probability that an event is an arrest is equal to the competing rates ratio $\mu/(\mu+\delta)$, which leaves the probability an event is career termination as $\delta/(\mu+\delta)$.

---

[32]At this point in the analysis, the impact of time spent incarcerated is not yet considered.

[33]For a more complete discussion of results associated with parallel processes and competing rates see Lawless (1982), pp. 484-491.

The following quantities are next derived using the parallel competing processes view of a criminal career:

- the probability an offender is still criminally active after t years,
- the probability an offender is re-arrested,
- the probability an offender is first re-arrested at time t,
- the distribution of time until an offender's next arrest conditioned on a next arrest occurring.

The probability an offender is still active after time t is dependent on the termination process and not on the arrest process. Since career length is distributed as an exponential, the probability a career lasts beyond t years is given by, $S(t) = e^{-\delta t}$.

The probability an offender is ever re-arrested is the same as the probability that the next event is an arrest, $\mu/(\mu+\delta)$, since if the next event is not an arrest then criminal activity is terminated. The probability an offender is first re-arrested at time t is equal to $\mu\, e^{-(\mu+\delta)t}\, dt$, which is derived as follows:

$$\Pr\begin{bmatrix}\text{RE-ARREST}\\ \text{AT TIME } t\\ \end{bmatrix} = \Pr\begin{bmatrix}\text{NO EVENT}\\ \text{OCCURS PRIOR}\\ \text{TO TIME } t\end{bmatrix} \times \Pr\begin{bmatrix}\text{RE-ARREST}\\ \text{AT TIME T}\\ \end{bmatrix}\begin{matrix}\text{NO EVENT}\\ \text{PRIOR TO}\\ \text{TIME } t\end{matrix}\Bigg]$$

$$= e^{-(\mu+\delta)t} \times \mu\, dt$$

<u>The distribution of time until an offender's next arrest conditioned on a next arrest occurring is exponential with parameter $(\mu+\delta)$</u>. This quantity derives from the probability that an arrest occurs at time t conditioned on the next event being an arrest:

$$\Pr\left[\begin{matrix}\text{NEXT ARREST}\\ \text{AT TIME T}\end{matrix}\,\middle|\,\begin{matrix}\text{NEXT EVENT}\\ \text{IS AN ARREST}\end{matrix}\right]$$

$$= \frac{\Pr[\text{NEXT ARREST AT TIME T}]}{\Pr[\text{NEXT EVENT IS AN ARREST}]}$$

$$= \frac{\Pr\left[\begin{matrix}\text{ARREST AT}\\ \text{TIME T}\end{matrix}\right] \times \Pr\left[\begin{matrix}\text{NO EVENT PRIOR}\\ \text{TO TIME T}\end{matrix}\right]}{\Pr[\text{NEXT EVENT IS AN ARREST}]}$$

$$= \frac{(\mu\ dt) \times e^{-(\mu+\delta)t}}{\frac{\mu}{\mu+\delta}}$$

$$= (\mu+\delta)\ e^{-(\mu+\delta)t}\ dt$$

Therefore, the conditional distribution of inter-arrest times for those who are re-arrested is exponential with parameter $(\mu+\delta)$.

Figure C-2 illustrates the effect of dropout on the distribution of inter-arrest times observed from a sample of offenders.

Curve A: + Distribution of inter-arrest times when no dropout occurs (μ = .2, δ = 0).

Curve B: • Degenerate distribution[34] of inter-arrest times after allowing career termination (μ = .2, δ = .1).

Curve C: * Distribution of inter-arrest times observed when μ=.2, δ =.1, conditional on a next arrest occurring (Curve B normalized so it has area equal to 1).

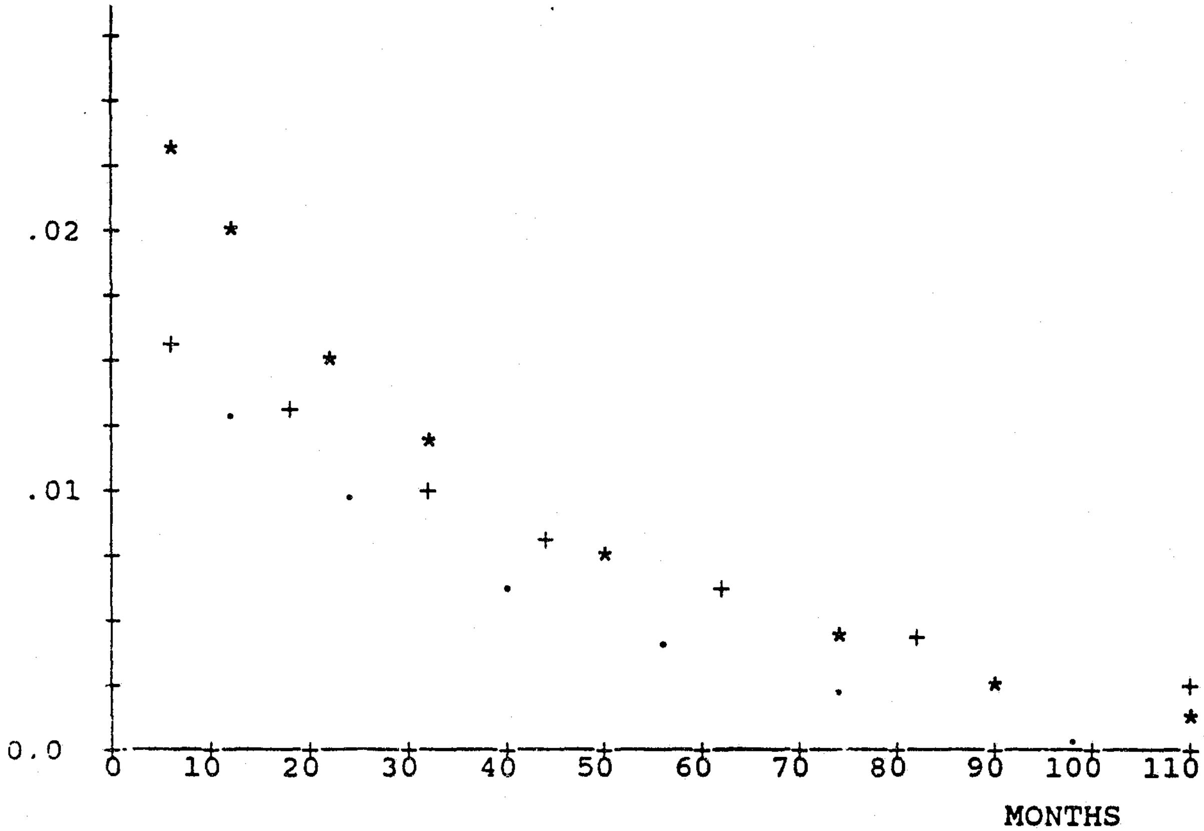


Figure C-2: The Effect of Career Termination of the Observed Rate of Arrest

[34]This probability distribution is degenerate in that the sum of the area under the curve is not equal to one.

A comparison between Curve A in Figure C-2 (the distribution of inter-arrest times without dropout) and Curve B (the probability of an arrest at time t with dropout) indicates that a larger proportion of long inter-arrest times are lost compared to short inter-arrest times as a result of a continuous termination process. This is because the probability that an offender is still criminally active, and thereby liable to be arrested, decreases as the length of time increases ($S(t) = e^{-\delta t}$).

Curve C shows this same effect of career termination on the conditional distribution of arrest intervals for those who are indeed re-arrested. After normalizing Curve B by the probability that an arrest does occur, Curve C reveals that the relative fraction of long intervals to a next arrests is smaller than in Curve A, while the relative fraction of short intervals is larger than in Curve A.

While, the continuous termination process does not alter the form of the inter-arrest time distribution, it does reduce the mean of the inter-arrest times. Even among sampled offenders who do not terminate offending before they are re-arrested, the average time until a next arrest reflects the influence of <u>both μ and δ</u>, rather than the effect of μ alone. Therefore, obtaining estimates of δ is important not only in characterizing the termination process; it also represents a necessary control for estimating the magnitude of the arrest rate.

The Likelihood Function

The maximum likelihood estimates of μ and δ for a group of offenders are the μ and δ which best explain the observed time each offender spent on the street, F, from the time of his target arrest until his next criterion arrest or through June, 1982, whichever occurred first. The maximum likelihood estimates are calculated by maximizing the sum of the log-likelihood of F over all offenders in a group:

$$\max_{\mu,\delta} \sum_{i=1}^{n} \text{Log}[L(F_i)]$$

Where, i = index of offenders in the group

The likelihood function for an individual allows for two possible outcomes, namely re-arrest is observed or not:

1) Criterion Re-Arrest Observed

```
    X<——J——><——————————F——————————>X

Target                              Next
Arrest                              Criterion
                                    Arrest
```

2) No Criterion Arrest Through June, 1982

```
    X<——J——><——————————————F——————————————>|

Target                                      6/82
Arrest
```

J = time spent incarcerated[35]
F = time spent free from prison

Figure C-3: Examples of Observations Used in Maximum Likelihood Estimates.

The likelihood function for Case 1 in which a criterion re-arrest is observed, is equal to the probability the offender does not dropout while incarcerated (a period of J months) multiplied

---

[35]In order to simplify the model, it is assumed that all time spent incarcerated occurs at the beginning of the follow-up interval. However, time spent incarcerated for non-criterion arrests which occur during the follow-up interval may occur toward the middle or end of the interval.

This simplification has no impact on the likelihood in the case of re-arrest observed. However, this simplification does have an impact when a re-arrest is not observed and the offender has a non-criterion arrest which leads to incarceration. In this case, the probability of not observing a criterion re-arrest is higher if the incarceration period occurs at the start of the follow-up interval as opposed to at the end. This simplification should have a small impact on the estimates of μ and δ because it affects only a few cases in the sample.

by the probability the offender is re-arrested after having been free a total of F months.

$$L(F,\text{Case}=1|J,\mu,\delta)$$

$$= \Pr\begin{bmatrix}\text{DOES NOT}\\ \text{DROPOUT}\\ \text{IN JAIL}\end{bmatrix} \times \Pr\begin{bmatrix}\text{RE-ARREST} & \text{DID NOT}\\ \text{F MONTHS} & \text{DROPOUT}\\ \text{LATER} & \text{IN JAIL}\end{bmatrix}$$

$$= e^{-\delta(J)} \times \mu\, e^{-(\mu+\delta)F}$$

$$= \mu\, e^{-\delta(J+F)} \times e^{-\mu(F)}$$

And the log-likelihood,

$$\text{Log}[L(\text{Case}=1)] = \log(\mu) - \delta(J+F) - \mu(F)$$

The likelihood function for case 2, re-arrest not observed, is equal to 1 minus the probability that an arrest is observed anytime within F.

$$L(F,\text{Case}=2|J,\mu,\delta) = 1 - \Pr\begin{bmatrix}\text{RE-ARREST}\\ \text{OBSERVED}\end{bmatrix}$$

$$= 1 - \Pr\begin{bmatrix}\text{DOES NOT}\\ \text{DROPOUT}\\ \text{IN JAIL}\end{bmatrix} \times \Pr\begin{bmatrix}\text{RE-ARREST} & \text{DOES NOT}\\ \text{WITHIN F} & \text{DROPOUT}\\ \text{MONTHS} & \text{IN JAIL}\end{bmatrix}$$

$$= 1 - \Pr\begin{bmatrix}\text{DOES NOT}\\ \text{DROPOUT}\\ \text{IN JAIL}\end{bmatrix} \times \Pr\begin{bmatrix}\text{NEXT EVENT} & \text{DOES NOT}\\ \text{WITHIN} & \text{DROPOUT}\\ \text{F MONTHS} & \text{IN JAIL}\end{bmatrix}$$

$$\times \Pr\begin{bmatrix}\text{NEXT} & \text{NEXT EVENT} & & \text{DOES NOT} \\ \text{EVENT IS} & \text{WITHIN} & \& & \text{DROPOUT} \\ \text{AN ARREST} & \text{F MONTHS} & & \text{IN JAIL}\end{bmatrix}$$

$$= 1 - \left[e^{-\delta J}\right] \times \left[1 - e^{-(\mu+\delta)F}\right] \times \left[\frac{\mu}{\mu + \delta}\right]$$

$$= 1 - \frac{\mu}{\mu + \delta} \; e^{-\delta J} + \frac{\mu}{\mu + \delta} \; e^{-\delta(J+F)} \; e^{-\mu(F)}$$

And the log-likelihood,

Log[L(Case=2)]

$$= \log\left[1 - \frac{\mu}{\mu + \delta} \; e^{-\delta J} + \frac{\mu}{\mu + \delta} \; e^{-\delta(J+F)} e^{-\mu(F)}\right]$$

Model Validation

The observed distribution of re-arrest times is used to assess the validity of the assumptions of the model -- δ and μ time-invariant and homogeneous across offenders. This test involves comparing the number and timing of re-arrests expected from the model to those actually observed in follow-up arrest data. Six re-arrest intervals are used: (1) offenders who are re-arrested within .5 years subsequent to release from jail, (2) those re-arrested from .5 to 1.0 years, (3) 1.0 to 1.5 years, (4) 1.5 to 2.5 years, (5) 2.5 to 3.5 years, (6) 3.5 to 5.0 years, and (7) 5.0 to 8.5 years.[36] The eighth and final category includes offenders not re-arrested within 8.5 years after release. The distribution of re-arrest times for each of the 44 offender groups in this study are analyzed in this manner.

Procedure:

Within any group, the expected number of offenders arrested in a particular interval is calculated as the sum over all offenders of each offender's *a priori* probability of being arrested in that interval.[37] The probability that an offender is arrested in a specific interval varies across individuals within a group

---

[36]These intervals were chosen to provide an expected minimum of 10 cases in each interval in a sample of 150 offenders for values of δ = .1 and μ = .2 per year.

[37]This probability does not depend on whether the offender actually was re-arrested.

based on the amount of time spent incarcerated subsequent to arrest and is calculated using the group's maximum likelihood parameter estimates, $\hat{\delta}$ and $\hat{\mu}$:

$$\Pr\begin{bmatrix}\text{OFFENDER}\\ \text{ARRESTED}\\ \text{BETWEEN}\\ (t_1,t_2)\end{bmatrix} = \int_{S=I_{1,i}}^{I_{2,i}} \Pr\begin{bmatrix}\text{A: DOES NOT}\\ \text{DROPOUT IN}\\ \text{JAIL}\end{bmatrix} \times \Pr\left[\begin{matrix}\text{B: NO}\\ \text{EVENT}\\ \text{PRIOR}\\ \text{TO S}\end{matrix}\middle| A\right]$$

$$\times \Pr\left[\begin{matrix}\text{C: ARREST}\\ \text{AT TIME}\\ \text{S}\end{matrix}\middle| A,B\right]$$

Where,

$$I_{1,i} = \begin{cases} t_1 & , \quad t_1 < G_i \\ G_i & , \quad t_1 \geq G_i \end{cases}$$

$$I_{2,i} = \begin{cases} t_2 & , \quad t_2 < G_i \\ G_i & , \quad t_2 \geq G_i \end{cases}$$

G_i = length of the time offender i is observed while free (time not in jail prior to June, 1982)

J_i = length of time spent in jail subsequent to target arrest.

$$= \int_{I_{1,i}}^{I_{2,i}} e^{-\delta J_i} \times e^{-(\mu+\delta)S} \times \mu \, dS$$

$$= \mu \, e^{-\delta J_i} \left[\frac{e^{-(\mu+\delta)S}}{-(\mu+\delta)} \right]_{I_{1,i}}^{I_{2,i}}$$

$$= \frac{\mu}{\mu+\delta} \, e^{-\delta J_i} \left[e^{-(\mu+\delta) I_{1,i}} - e^{-(\mu+\delta) I_{2,i}} \right]$$

For the final category, the probability an offender is not re-arrested within 5 years is as follows:

$$\Pr\begin{bmatrix} \text{NO} \\ \text{RE-ARREST} \\ \text{BEFORE} \\ \text{FIVE YEARS} \end{bmatrix} = 1 - \int_0^{I_{5,i}} \Pr\begin{bmatrix} \text{RE-ARREST} \\ \text{AT TIME S} \end{bmatrix}$$

Where,

$$I_{5,i} = \begin{cases} 5 & , \quad 5 < G_i \\ G_i & , \quad 5 \geq G_i \end{cases}$$

$$= 1 - \int_0^{I_{5,i}} e^{-\delta J_i} \, e^{-(\mu+\delta)S} \mu \, dS$$

$$= 1 - \frac{\mu}{\mu+\delta} \, e^{-\delta J_i} \left[1 - e^{-(\mu+\delta) \times I_{5,i}} \right]$$

Pearson's χ^2 statistic is used to determine the model's goodness of fit:

$$\text{Pearson's Chi-Square} = \sum_{b=1}^{7} \frac{\left(O_b - E_b\right)^2}{E_b} \sim \chi^2 \text{ (with 6 d.o.f)}$$

where,

b = index of intervals

O_b = observed number of offenders in b

E_b = expected number of offenders in b

Pearson's χ^2 statistic asymptotically has a chi-square distribution with 6 degrees of freedom as the number of offenders in a group gets large.

Result:

Table C-1 reports the results of the analysis for the 44 groups in this study. Under the hypothesized model, it is expected, at the $\alpha=.01$ level of significance, that the chi-square statistic will be less than 18.48 for any group. The first two groups fail to meet this criterion; however, the sample size of these groups is quite large (over 1000 offenders) which greatly increases the ability of the test to detect slight variations between expected and observed re-arrest time distributions, and so the test failure in these two groups can reasonably be attributed to the excessively high power of the χ^2 test in large samples.

Of the remaining 42 groups 39 meet the $\alpha=.01$ level test criterion which shows generally good conformance between the observed data and the assumed model.[38] Therefore, it can be concluded that the model assumptions of δ and μ time-invariant and homogeneous across offenders within a group is consistent with the observed data.

[38]The probability that 3 or fewer tests out of 42 will fail at the α=.01 level is .9992. The probability of 2 or fewer failures is .9914.

Table C-1: Results of Model Validation by Group

Rank	Attribute ID	Sample Size	χ^2 (6 d.f.)
1	W-11-00-000	2700	72.28**
2	B-11-00-000	1151	30.51**
3	W-11-00-00D	888	7.71
4	W-22-00-V00	790	7.71
5	B-11-00-0R0	699	4.75
6	W-22-00-000	696	6.67
7	B-22-00-000	534	5.89
8	W-**-00-0R0	483	16.98*
9	B-11-00-V00	445	10.74
10	W-33-00-*00	428	5.91
11	B-22-00-V00	398	10.65
12	W-11-00-00D	397	17.01*
13	B-22-00-0R0	362	31.67**
14	W-34-00-*00	315	2.84
15	W-22-00-00D	312	14.03
16	B-**-10-000	286	4.63
17	W-**-10-000	291	17.21*
18	B-**-1I-0R0	259	14.03
19	W-**-10-V00	261	4.10
20	B-**-10-V*0	254	16.73*
21	B-12-3I-V*0	228	13.17
22	B-*2-1I-V*0	212	13.78
23	W-22-00-V0D	230	7.42
24	B-**-10-0R0	201	22.53**
25	W-22-00-0R0	203	9.67
26	B-**-1I-000	175	3.67
27	B-12-3I-000	164	6.98
28	W-11-00-0RD	173	2.50
29	B-**-00-00D	169	3.14
30	W-12-3I-0*0	160	9.59
31	W-*3-**-V*0	167	4.90
32	W-**-10-00D	156	11.71
33	B-**-*0-0*D	151	9.57
34	B-33-00-*00	162	1.85
35	B-12-3I-010	146	22.31**
36	W-11-00-V0D	151	8.14
37	W-12-3I-V*0	142	11.03
38	B-34-00-V00	136	2.15
39	W-**-10-**D	128	5.89
40	B-12-3I-**D	120	5.48
41	B-22-3I-**0	116	14.61*
42	B-*3-3I-V*0	106	11.93
43	W-**-1I-*00	105	4.23
44	W-**-10-0R0	80	6.41

* statistically significant at the α=.05 level
** statistically significant at the α=.01 level

An Infinite-Horizon Maximum-Likelihood Estimator

Under the following ideal conditions, the maximum-likelihood estimators for μ and δ have closed-form solutions:

1. The parameters μ and δ are time-invariant over the follow-up period for each individual.

2. The parameters μ and δ are homogeneous across all offenders included in a group.

3. The follow-up study period is infinitely long, which allows enough time for every offender to be re-arrested if they are re-arrested.

4. Offender's do not spend any time incarcerated.

The addition of assumptions 3 and 4 greatly simplify the likelihood function, particularly for the case in which a re-arrest is not observed.

The maximum likelihood estimators for μ and δ are found by taking the derivative of the log-likelihood of the follow-up intervals observed for a sample and setting it equal to zero. The derivative of the log-likelihood for the sample is equal to the sum of the derivatives of the log-likelihood for each individual.

The likelihood an individual offender is arrested after F months is as follows:

$$L_i = \Pr\left[\begin{matrix}\text{ARREST AT}\\ \text{TIME } F_i\end{matrix}\middle|\begin{matrix}\text{NO PRIOR}\\ \text{EVENT}\end{matrix}\right] \times \Pr\left[\begin{matrix}\text{NO EVENT}\\ \text{PRIOR TO } F_i\end{matrix}\right]$$

$$= \mu \, dt \, e^{-(\mu+\delta)F_i}$$

and the log-likelihood,

$$\text{Log}[L_i] = \log(\mu) + \log(dt) - (\mu+\delta)F_i$$

and the derivatives of the log-likelihood,

$$\frac{d(\text{Log-}L_i)}{d\mu} = \frac{1}{\mu} - F_i$$

$$\frac{d(\text{Log-}L_i)}{d\delta} = - F_i \quad .$$

The likelihood an individual offender is never re-arrested is the probability career termination precedes the next arrest that would have occurred:

$$L_i = \Pr\begin{bmatrix}\text{DROPOUT OCCURS}\\ \text{BEFORE THE}\\ \text{NEXT ARREST}\end{bmatrix} = \frac{\delta}{\mu + \delta}$$

and the log-likelihood,

$$\text{Log-}L_i = \log\left[\frac{\delta}{\mu + \delta}\right]$$

$$= \log(\delta) - \log(\mu + \delta)$$

and the derivatives of the log-likelihood,

$$\frac{d(\text{Log-}L_i)}{d\mu} = \frac{-1}{\mu + \delta}$$

$$\frac{d(\text{Log-}L_i)}{d\delta} = \frac{1}{\delta} - \frac{1}{\mu + \delta} \quad .$$

So, the sum of the individual derivatives of the log-likelihood with respect to μ is the following:

$$\frac{d(\text{Log-L})}{d\mu} = \sum_{i} \frac{d(\text{Log-L}_i)}{d\mu}$$

$$= \sum_{\substack{\text{all offenders} \\ \text{who are} \\ \text{re-arrested}}} \frac{d(\text{Log-L}_i)}{d\mu} + \sum_{\substack{\text{all offenders} \\ \text{who are not} \\ \text{re-arrested}}} \frac{d(\text{Log-L}_i)}{d\mu}$$

$$= \Sigma \left[\frac{1}{\mu} - F_i\right] + \Sigma \left[\frac{-1}{\mu + \delta}\right]$$

$$= \frac{(1-P) \times N}{\mu} - F \times (1-P) \times N - \frac{P \times N}{\mu + \delta}$$

Where,

N = the sample size

P = proportion of offenders not re-arrested

F = average re-arrest time for those offenders who are re-arrested

And, the sum of the individual derivatives of the log-likelihood with respect to δ is the following:

$$\frac{d(\text{Log-L})}{d\delta} = \sum_{i} \frac{d(\text{Log-L}_i)}{d\delta}$$

$$= \sum_{\substack{\text{all offenders} \\ \text{who are} \\ \text{re-arrested}}} \frac{d(\text{Log-L}_i)}{d\delta} + \sum_{\substack{\text{all offenders} \\ \text{who are not} \\ \text{re-arrested}}} \frac{d(\text{Log-L}_i)}{d\delta}$$

$$= \Sigma \left[- F_i\right] + \Sigma \left[\frac{1}{\delta} - \frac{1}{\mu + \delta}\right]$$

$$= -F \times N \times (1-P) + \frac{P \times N}{\delta} - \frac{P \times N}{\mu + \delta}$$

Setting the derivatives equal to zero yields two equations with two unknowns, μ and δ:

1) $$\frac{(1-P) \times N}{\mu} - F \times N \times (1-P) - \frac{P \times N}{\mu + \delta} = 0 \quad \text{(C-}$$

$$-F \times N \times (1-P) + \frac{P \times N}{\delta} - \frac{P \times N}{\mu + \delta} = 0$$

Which imply,

$$\frac{(1-P)}{\mu} - \frac{P}{\mu + \delta} = (1-P) \times F$$

$$\frac{P}{\delta} - \frac{P}{\mu + \delta} = (1-P) \times F$$

From which the following maximum-likelihood estimators for μ and δ are obtained:

$$\mu = \frac{(1-P)}{F} \quad , \quad \delta = \frac{P}{F}$$

A Two-Population Process

If the data in an infinite horizon model comes from a sample of offenders which contains two sub-populations on which the one-population, maximum-likelihood estimators are used, the following results are obtained:

$$\mu = \frac{\Pr\left[\text{RE-ARREST}\right]}{E\begin{bmatrix}\text{INTER-ARREST} \\ \text{TIME FOR THOSE} \\ \text{WHO ARE RE-ARRESTED}\end{bmatrix}}$$

and,

$$\Pr\left[\text{RE-ARREST}\right] = \Pr\left[\text{RE-ARREST}\,\middle|\,\begin{matrix}\text{FROM} \\ \text{GROUP 1}\end{matrix}\right] \times \Pr\begin{bmatrix}\text{FROM} \\ \text{GROUP 1}\end{bmatrix}$$

$$+ \Pr\left[\text{RE-ARREST}\,\middle|\,\begin{matrix}\text{FROM} \\ \text{GROUP 2}\end{matrix}\right] \times \Pr\begin{bmatrix}\text{FROM} \\ \text{GROUP 2}\end{bmatrix}$$

$$= \frac{\mu_1}{\mu_1 + \delta_1} \times Q + \frac{\mu_2}{\mu_2 + \delta_2} \times (1-Q)$$

where, Q = proportion of offenders in group 1.

$$E\begin{bmatrix}\text{INTER-ARREST} \\ \text{TIME FOR THOSE} \\ \text{WHO ARE RE-ARRESTED}\end{bmatrix}$$

$$= E\left[\begin{matrix}\text{INTER-ARREST} \\ \text{TIME FOR THOSE} \\ \text{WHO ARE RE-ARRESTED}\end{matrix}\,\middle|\,\begin{matrix}\text{RE-ARRESTEE} \\ \text{IS FROM} \\ \text{GROUP 1}\end{matrix}\right] \times \Pr\begin{bmatrix}\text{RE-ARRESTEE} \\ \text{IS FROM} \\ \text{GROUP 1}\end{bmatrix}$$

$$+ E\begin{bmatrix}\text{INTER-ARREST} \\ \text{TIME FOR THOSE} \\ \text{WHO ARE RE-ARRESTED}\end{bmatrix}\begin{matrix}\text{RE-ARRESTEE} \\ \text{IS FROM} \\ \text{GROUP 2}\end{bmatrix} \times Pr\begin{bmatrix}\text{RE-ARRESTEE} \\ \text{IS FROM} \\ \text{GROUP 2}\end{bmatrix}$$

$$E\left[\begin{matrix}\text{INTER-ARREST} \\ \text{TIME FOR THOSE} \\ \text{WHO ARE RE-ARRESTED}\end{matrix}\middle|\begin{matrix}\text{RE-ARRESTEE} \\ \text{IS FROM} \\ \text{GROUP 1}\end{matrix}\right] = \frac{1}{\mu_1 + \delta_1}$$

$$Pr\begin{bmatrix}\text{RE-ARRESTEE} \\ \text{IS FROM} \\ \text{GROUP 1}\end{bmatrix} = \frac{Q \times \dfrac{\mu_1}{\mu_1 + \delta_1}}{Q \times \dfrac{\mu_1}{\mu_1 + \delta_1} + (1-Q) \times \dfrac{\mu_2}{\mu_2 + \delta_2}}$$

And the quantities are analogously found for group 2, and, so, the maximum likelihood estimate for μ is as follows:

$$\mu = \frac{\left[\dfrac{Q \times \mu_1}{\mu_1 + \delta_1} + \dfrac{(1-Q) \times \mu_2}{\mu_2 + \delta_2}\right]}{\left[\dfrac{\left[\dfrac{Q \times \mu_1}{(\mu_1 + \delta_1)^2} + \dfrac{(1-Q) \times \mu_2}{(\mu_2 + \delta_2)^2}\right]}{\left[\dfrac{Q \times \mu_1}{\mu_1 + \delta_1} + \dfrac{(1-Q) \times \mu_2}{\mu_2 + \delta_2}\right]}\right]} \qquad \text{(C-2)}$$

And analogously for δ:

$$\delta = \frac{1 - Pr[\text{RE-ARREST}]}{E\begin{bmatrix}\text{INTER-ARREST} \\ \text{TIME FOR THOSE} \\ \text{WHO ARE RE-ARRESTED}\end{bmatrix}}$$

$$\delta = \frac{\left[\dfrac{Q \times \delta_1}{\mu_1 + \delta_1} + \dfrac{(1-Q) \times \delta_2}{\mu_2 + \delta_2}\right]}{\left[\dfrac{\left[\dfrac{Q \times \mu_1}{(\mu_1 + \delta_1)^2} + \dfrac{(1-Q) \times \mu_2}{(\mu_2 + \delta_2)^2}\right]}{\left[\dfrac{Q \times \mu_1}{\mu_1 + \delta_1} + \dfrac{(1-Q) \times \mu_2}{\mu_2 + \delta_2}\right]}\right]} \tag{C-3}$$

Appendix D: Simulation Results

The small sample properties of the maximum likelihood estimator are tested under controlled conditions using simulated offender follow-up arrest data. Whereas maximum likelihood provides asymptotically consistent estimates, for small samples maximum likelihood estimates tend to be biased. This analysis is used to establish the expected bias and the coefficient of variation (c.o.v.) -- the standard error divided by the mean-- for $\hat{\delta}$ and $\hat{\mu}$ as the underlying behavioral parameters (δ and μ) vary, in addition variation in both sample size and mean length of offender follow-up observation period.

Data Generator:

The data generator creates a group of simulated offender follow-up arrest records similar to those observed in the Detroit dataset. The input parameters to the data generator include the following: 1) δ - the termination rate, 2) μ - the arrest rate, 3) C - the average follow-up observation period, and 4) n - the sample size. The simulator returns "n" offender follow-up records describing how much time, if any, each offender spent in jail, and how long until either the offender is re-arrested or the follow-up observation period ends.

The data generator uses a pseudo-random number generator to determine each offender's simulated follow-up period. The probability an offender is sent to jail is 50%. For those offenders who do go to jail, the time served is drawn from an exponential distribution with a mean time served of 1 year.

Each offender is randomly assigned a date of next arrest, career termination, and end of follow-up period. The time between the date of release from prison and the next arrest, consistent with the Poisson arrest process assumed in the model, is drawn from an exponential distribution with a mean of $1/\mu$. Similarly, the time between the target arrest and career termination is drawn from an exponential distribution with mean of $1/\delta$. The length of time from the target arrest to the end of the follow-up observation period is drawn from a uniform distribution of width 4 years with a mean of C. For the Detroit data, the observation period ends in June 1982 and the length of an offender's observation period varies from 4.5 to 8.5 years with an approximate mean of C=6.5 years depending on the date of the target arrest. If the date of the next arrest occurs before both the termination date and the end of the observation period date, then the offender is reported to have been re-arrested. Otherwise, the data generator reports that no re-arrest was observed.

The Analysis:

For each set of input parameters, 30 groups of simulated offender data are created (the same 30 random number generator seeds are used to generate the data for each set). The range of input parameter values was chosen to simulate the range of values observed in the Detroit and New York City databases. The behavioral parameter values (δ,μ) include (.1,.2), the average values for Detroit, and (.2,.4), the New York City average, and other similar combinations of (δ,μ). The largest mean length of follow-up horizon explored is 6.5 years, the average follow-up observation period for the Detroit data. To explore the effect of a shorter horizon, the values 2.0, 3.0, and 4.5 years are also considered. The sample sizes explored range from a small group of just 150 offenders up to a large group of 1500 offenders.

The maximum likelihood estimator is then used to estimate the parameters (δ,μ) for each group of offenders. Table D-1 reports for each set of input parameter values the average bias (e.g.,$(\hat{\delta} - \delta)/\delta$) and the coefficient of variation across the 30 simulated groups (e.g., $s(\hat{\delta})/E(\hat{\delta})$) for both $\hat{\delta}$ and $\hat{\mu}$, respectively.[39]

[39]Whereas the MLE can provide an estimate of the c.o.v. for $\hat{\delta}$ and $\hat{\mu}$, this estimate may be highly inaccurate for small samples. Therefore, the sample coefficient of variation $(s(\hat{\delta})/e(\hat{\delta}))$ is used to examine the variation in estimates obtained over the 30 simulated groups generated for each set of parameter values.

Table D-1: Bias and C.O.V. for Simulated Offender Groups

	δ	μ	C	n	$\hat{\delta}$: Bias	c.o.v.	$\hat{\mu}$: Bias	c.o.v.
1	0.05	0.1	6.5	150	0.126	0.883	0.019	0.211
2	0.05	0.2	6.5	150	-0.050	0.592	-0.022	0.155
3	0.05	0.4	6.5	150	0.051	0.421	-0.017	0.126
4	0.1	0.1	6.5	150	0.072	0.575	0.021	0.221
5	0.1	0.2	2.0	150	0.852	0.571	0.146	0.219
6	0.1	0.2	2.0	500	0.668	0.424	0.084	0.116
7	0.1	0.2	2.0	1500	0.577	0.230	0.077	0.068
8	0.1	0.2	3.0	150	0.489	0.697	0.100	0.193
9	0.1	0.2	3.0	500	0.430	0.389	0.072	0.103
10	0.1	0.2	3.0	1500	0.340	0.270	0.063	0.075

Table D-1: Continued

	δ	μ	C	n	$\hat{\delta}$: Bias	c.o.v.	$\hat{\mu}$: Bias	c.o.v.
11	0.1	0.2	4.5	150	0.295	0.372	0.081	0.150
12	0.1	0.2	4.5	500	0.254	0.270	0.073	0.123
13	0.1	0.2	4.5	1500	0.308	0.182	0.087	0.080
14	0.1	0.2	6.5	150	-0.026	0.375	-0.027	0.148
15	0.1	0.2	6.5	500	0.110	0.206	0.038	0.100
16	0.1	0.2	6.5	1500	0.066	0.150	0.020	0.071
17	0.1	0.3	2.0	150	0.562	0.876	0.116	0.210
18	0.1	0.3	2.0	500	0.380	0.560	0.052	0.137
19	0.1	0.3	2.0	1500	0.261	0.338	0.023	0.067
20	0.1	0.3	3.0	150	0.157	0.609	0.047	0.166
21	0.1	0.3	4.5	150	0.303	0.362	0.073	0.171
22	0.1	0.3	6.5	150	0.086	0.310	0.023	0.140
23	0.1	0.4	2.0	150	0.220	0.790	0.054	0.195
24	0.1	0.4	2.0	500	0.354	0.473	0.052	0.146
25	0.1	0.4	2.0	1500	0.270	0.298	0.034	0.068
26	0.1	0.4	3.0	150	0.118	0.590	0.029	0.164
27	0.1	0.4	4.5	150	0.084	0.372	0.034	0.140
28	0.1	0.4	6.5	150	0.193	0.286	0.058	0.118
29	0.2	0.1	6.5	150	0.041	0.472	0.048	0.310
30	0.2	0.2	2.0	150	0.291	0.622	0.092	0.268
31	0.2	0.2	2.0	500	0.221	0.353	0.035	0.150
32	0.2	0.2	2.0	1500	0.072	0.190	0.000	0.078
33	0.2	0.2	3.0	150	0.192	0.570	0.071	0.228
34	0.2	0.2	3.0	500	0.078	0.273	0.006	0.136
35	0.2	0.2	3.0	1500	0.047	0.131	0.005	0.061
36	0.2	0.2	4.5	150	0.071	0.353	0.030	0.197
37	0.2	0.2	4.5	500	0.027	0.147	0.002	0.098
38	0.2	0.2	4.5	1500	0.035	0.091	0.004	0.062
39	0.2	0.2	6.5	150	-0.021	0.265	-0.014	0.185

40	0.2	0.2	6.5	500	0.003	0.113	-0.014	0.091
41	0.2	0.2	6.5	1500	-0.006	0.080	-0.015	0.065
42	0.2	0.3	2.0	150	0.184	0.612	0.069	0.226
43	0.2	0.3	2.0	500	0.072	0.357	0.012	0.133
44	0.2	0.3	2.0	1500	0.028	0.197	-0.007	0.064
45	0.2	0.3	3.0	150	0.056	0.415	0.018	0.189
46	0.2	0.3	4.5	150	0.008	0.275	-0.014	0.173
47	0.2	0.3	6.5	150	0.007	0.242	-0.007	0.160
48	0.2	0.4	2.0	150	0.047	0.475	0.024	0.185
49	0.2	0.4	2.0	500	0.068	0.235	0.009	0.098

Table D-1: Continued

	δ	μ	c	n	$\hat{\delta}$: Bias	c.o.v.	$\hat{\mu}$: Bias	c.o.v.
50	0.2	0.4	2.0	1500	0.044	0.[illegible]07	0.000	0.057
51	0.2	0.4	3.0	150	0.048	0.384	0.010	0.174
52	0.2	0.4	3.0	500	0.021	0.160	-0.014	0.091
53	0.2	0.4	3.0	1500	0.023	0.124	-0.008	0.053
54	0.2	0.4	4.5	150	-0.008	0.297	-0.015	0.136
55	0.2	0.4	4.5	500	-0.012	0.131	-0.025	0.082
56	0.2	0.4	4.5	1500	-0.007	0.069	-0.017	0.054
57	0.2	0.4	6.5	150	-0.012	0.246	-0.020	0.155
58	0.2	0.4	6.5	500	-0.011	0.123	-0.030	0.074
59	0.2	0.4	6.5	1500	-0.002	0.051	-0.016	0.044
60	0.3	0.1	2.0	150	0.295	0.701	0.182	0.334
61	0.3	0.1	2.0	500	0.131	0.432	0.029	0.195
62	0.3	0.1	2.0	1500	0.011	0.278	-0.010	0.108
63	0.3	0.1	3.0	150	0.215	0.502	0.162	0.350
64	0.3	0.1	4.5	150	0.201	0.415	0.127	0.273
65	0.3	0.1	6.5	150	0.095	0.339	0.087	0.279
66	0.3	0.2	6.5	150	0.001	0.233	0.015	0.206
67	0.3	0.4	6.5	150	-0.018	0.212	-0.018	0.169

Summary of Simulation Findings:

The relationships of bias and c.o.v. of $\hat{\delta}$ and $\hat{\mu}$ as a function of input parameters are each summarized using a multiplicative sensitivity model such as the following[40]:

$$|\text{Bias}(\hat{\delta})| = e^{\beta_0} \delta^{\beta_1} \mu^{\beta_2} C^{\beta_3} n^{\beta_4}$$

In these sensitivity models, a negative coefficient indicates an inverse relationship such that increases in the parameter effect a decrease in the dependent variable resulting in greater accuracy (smaller bias) or greater precision (smaller c.o.v.) for estimates associated with larger values of the parameter. At a coefficient value of -1, the magnitude of the effect is identical (e.g., a doubling in the parameter results in halving the dependent variable). For larger negative coefficients, the parameter has a strong influence on the dependent variable. For example, a coefficient of -2 indicates a quadratic effect such that if the independent parameter value were to be doubled, the dependent variable value would be reduced to one fourth its original size. For negative coefficients closer to zero, the decrease in the dependent variable is smaller. Table D-2 reports the results of this summary analysis.

[40]The coefficients of this model are estimated using least-squares regression after taking the logarithms of both independent and dependent variables. For most sets of input parameters a positive bias is observed. However, a few sets exhibit a small negative bias. To avoid taking a logaraithm of a negative number, absolute bias is used instead of bias in these summary equations.

Table D-2: Sensitivity Analysis of Bias and C.O.V of $\hat{\delta}$ and $\hat{\mu}$ With Respect to Input Parameters[41]

Coeff. for	Bias($\hat{\delta}$)	c.o.v.($\hat{\delta}$)	Bias($\hat{\mu}$)	c.o.v.($\hat{\mu}$)
Interc	-3.12 (.96)	1.15 (.20)	-2.20 (1.40)^{ns}	.71 (.13)
δ	-1.78 (.22)	- .52 (.05)	- .76 (.33)	.09 (.03)
μ	-1.09 (.23)	- .42 (.05)	- .71 (.33)	- .40 (.03)
C	-2.00 (.22)	- .77 (.05)	- .41 (.32)^{ns}	- .26 (.03)
n	- .32 (.11)	- .50 (.02)	- .59 (.16)	- .48 (.02)
R^2	.70	.92	.28	.95
Detroit	.073	.390	.050	.174
NY City	.010	.203	.018	.141

Overall, the pervasiveness of negative coefficient indicates that a decrease in any parameter, δ, μ, C or n, results in a loss of information for use in estimation. The nature of the loss is quite clear when sample size or observation period is reduced. For the behavioral parameters, a decrease in either δ or μ means that fewer events (i.e., arrests and career terminiations) are expected to be observed within a fixed period of time. Thus, the amount of information available from a sample is also reduced. When less information is available the quality of the estimates $\hat{\delta}$ and $\hat{\mu}$ decreases and so the bias and c.o.v. increase. All the coefficients reported in Table D-2 are negative except one. The exception to this pattern in the c.o.v.($\hat{\mu}$) which decreases slightly as δ decreases. This decrease occurs since as δ gets smaller, the

[41]The standard error of each coefficient is reported in parentheses. All coefficients are significantly different from zero in a two-tailed test at the α=.05 level of significance except the two marked by "ns."

re-arrest data becomes more and more indicative of μ due to fewer career terminations.

In general, the accuracy of δ estimates reflected in bias($\hat{\delta}$) is most sensitive to any parameter changes. A reduction in the horizon length has the largest impact. A decrease of average follow-up observation time from 6.5 to 2.0 years, roughly a 3-fold reduction, results in a 10-fold increase in bias($\hat{\delta}$) - $(2.0/6.5)^{-2.0}$ = 10.6. In contrast, the benefits of increased sample size due not accrue very quickly. An increase in the sample size from 150 to 1500 offenders, a 10-fold increase, results in only a 2-fold reduction in bias($\hat{\delta}$) -- $(1500/150)^{-.32}$ = .48.

The accuracy of estimates of μ (i.e., bias($\hat{\mu}$)) is much less sensitive to similar parameter changes than estimates of δ. However, parameter changes have very similar effects on the precision of both δ and μ, reflected similar sensitivity coefficients for c.o.v.($\hat{\delta}$) and c.o.v.($\hat{\mu}$).

The bias and coefficient of variation predicted by the sensitivity model for two sets of input parameters are also reported in Table D-2. The first, referred to as Detroit, uses average values of δ and μ observed across offender groups in Detroit, the average horizon length and the minimum sample size for the Detroit data: $(\delta,\mu,c,n) = (.1,.2,6.5,150)$. The second set, referred to as New York City, uses the same average horizon length and sample size and the average values of δ and μ observed across offender groups in the New York City data: $(\delta,\mu,c,n) = (.2,.4,6.5,150)$. The biases in $\hat{\delta}$ and $\hat{\mu}$ expected using the Detroit parameters are under 10% (.07 and .05, respectively). With the increases in δ and μ, these biases are substantially reduced for the New York City data to .01 and .02, respectively.

Simulation is used to establish a target minimum, within group sample-size, and to estimate the impact of heterogeneity on the maximum likelihood estimates.

Minimum Target Sample Size

Simulated data are used to test the maximum likelihood estimation procedure under controlled conditions in order to determine a minimum sample size. This minimum sample size is determined as the smallest sample for which an acceptable variance for the estimate of δ is obtained.

A pseudo-random number generator is used to create simulated offender records in samples of varying size, using the same parameter values $(\mu,\delta) = (.5,.1)$ in each sample. The results of estimating μ and δ by sample size are reported in Table D-3 . The first column indicates the number of samples of a given size explored. The standard errors reported are an average across samples of the same size of the standard errors of the estimates obtained.

Table D-3 also reports the coefficient of variation (C.V.) which is defined as the standard error divided by the expected value. In this case, the expected value of μ is .5 and δ is .1, the values used to generate the data. The C.V. is a scale independent measure of the accuracy of an estimate and is useful for comparing the accuracy with which μ is calculated to the accuracy with which δ is calculated.

# of Samples	n	$S.E._\mu$	$C.V._\mu$	$S.E._\delta$	$C.V._\delta$
5	40	.1140	.22	.0469	.46
4	80	.0839	.16	.0334	.33
4	160	.0573	.11	.0214	.21
2	320	.0396	.07	.0152	.15
1	640	.0285	.05	.0111	.11

Table D-3: Accuracy of the One-Population Maximum Likelihood Estimate as Sample Size Varies

Table D-3 reveals that although the standard error of δ is less than that of μ, the C.V. for δ is larger than the C.V. for μ at every sample size analyzed. In fact, μ is estimated about twice as accurately as δ for every sample size.

Table D-3 also shows how the standard error of the estimate of δ varies with the sample-size under simulated conditions. These standard errors obtained under controlled conditions are used to determine the target minimum sample-size. An error in the estimate of δ of .05 is chosen as a "reasonably large" error. Therefore, it is desirable for the likely error in the estimate of δ to be less than .05. Assuming the error comes from a normal distribution, then if the standard error $S.E.(\delta) = .025$, 95% of the time a confidence interval of the maximum likelihood estimate of δ +/- .05 should contain the true value of the parameter δ ($S.E. = .05/1.96 \approx .025$). Table D-3 shows a standard error of .025 corresponds to a sample size of approximately 150 offenders per group. Thus, the target minimum sample size for the maximum likelihood estimates is 150 offenders per group.

Effect of Heterogeneity

The effect of actual population heterogeneity within a group on the estimates derived assuming homogeneous μ and δ was tested by submitting data derived from two distinct sub-populations to the maximum likelihood code for a single population. The two population data is characterized by five parameters: the arrest and termination rates for sub-population 1 μ_1 and δ_1, the arrest and termination rates for sub-population 2 μ_2 and δ_2, and a probability Q that an offender belongs to sub-population 1.[42]

Table D-4 displays estimates of homogeneous (one population) μ and δ obtained by varying the parameters used to generate the simulated data for the two sub-populations. The results are very similar to the values predicted by the infinite horizon likelihood estimator (Appendix C, equations C-2 and C-3). These results are also very similar to a weighted average of the parameters, $[(Q\times\mu_1 + (1-Q)\times\mu_2),(Q\times\delta_1 + (1-Q)\times\delta_2)]$ and a weighted average of the inverse of the parameters, $[(Q/\mu_1 + (1-Q)/\mu_2)^{-1},(Q/\delta_1 + (1-Q)/\delta_2)^{-1}]$. On the whole, the one-population maximum likelihood estimates of (μ,δ) provide estimates that are akin to average values of μ and δ for all offenders in the groups even when two distinct sub-populations

---

[42]The likelihood model used in this analysis assumes the data come from a single population, while in this case the data actually come from two sub-populations. A likelihood function which estimates the parameters of the two sub-populations was also analyzed. In general, it was found that a sample size of 160 simulated offenders is too small to accurately estimate the five parameters of a two-population model.

are present.[43]

[43] In one clear exception to this pattern in the table, the weighted average of the parameter (AVG) is seen to be highly sensitive to changes in parameters in the smaller sub-population. None of the other estimates display this same sensitivity over the range of parameters explored.

Parameter Values Used to Generate Simulated Data					Values for δ:				Values for μ:			
μ_1	δ_1	μ_2	δ_2	Q	EST	INF	AVG	INV	EST	INF	AVG	INV
.5	.1	.5	.1	.8	.10	.10	.10	.10	.50	.50	.50	.50
.5	.1	1.0	.1	.8	.09	.10	.10	.10	.55	.56	.60	.56
.5	.1	1.0	.3	.8	.11	.12	.14	.12	.54	.55	.60	.56
.5	.1	1.0	.5	.8	.13	.13	.18	.12	.51	.53	.60	.56
.5	.1	1.0	1.0	.8	.14	.15	.28	.12	.49	.51	.60	.56
.5	.1	1.0	2.0	.8	.15	.17	.48	.12	.45	.47	.60	.56
.5	.1	.5	.5	.8	.13	.15	.18	.12	.48	.49	.50	.50
.5	.1	2.0	.5	.8	.12	.12	.18	.12	.57	.58	.80	.59
.5	.1	4.0	.5	.8	.12	.11	.18	.12	.60	.62	1.20	.61
.5	.1	1.0	.5	.2	.36	.33	.42	.28	.78	.77	.90	.83
.5	.1	1.0	.5	.5	.22	.20	.30	.17	.63	.61	.75	.67
.5	.1	1.0	.5	.9	.11	.12	.14	.11	.50	.52	.55	.53
.5	.1	1.0	.5	.95	.10	.11	.12	.10	.50	.51	.53	.51
.5	.05	1.0	.1	.8	.05	.06	.06	.06	.53	.56	.60	.56
.5	.1	1.0	.1	.8	.09	.10	.10	.10	.55	.56	.60	.56
.5	.3	1.0	.1	.8	.27	.27	.26	.21	.60	.59	.60	.56
.5	.5	1.0	.1	.8	.38	.43	.42	.28	.60	.60	.60	.56
.5	.05	1.0	.5	.8	.08	.09	.14	.06	.50	.52	.60	.56
.5	.1	1.0	.5	.8	.12	.13	.18	.12	.51	.53	.60	.56
.5	.3	1.0	.5	.8	.34	.33	.34	.33	.58	.56	.60	.56
.5	.5	1.0	.5	.8	.48	.51	.50	.50	.58	.58	.60	.56

EST: Result obtained by maximum likelihood estimation.

INF: Value predicted by infinite horizon maximum likelihood formula.

AVG: Weighted average of the two groups parameter values.

INV: Inverse of the weighted average of the inverse of the two groups parameter values.

Table D-4: Maximum Likelihood Estimates Obtained with a One-Population Model and Two Sub-Population Data

These results show that the one-population likelihood function estimates of (μ,δ) can be considered to be a type of average values of μ and δ for all offenders in the group even when two distinct sub-populations are present.

Appendix E: Results of Maximum Likelihood estimation.

RANK	ATTRIBUTE ID	SAMPLE SIZE	MAXIMUM LIKELIHOOD ESTIMATE OF: μ	δ	STANDARD DEV. OF MAX. LIKELIHOOD ESTIMATE OF: μ	δ
1	W-11-00-000	2700	.2583	.1909	.0094	.0099
2	B-11-00-000	1151	.3128	.0479	.0146	.0081
3	W-11-00-00D	888	.1763	.1654	.0135	.0209
4	W-22-00-V00	790	.0757	.1061	.0087	.0309
5	B-11-00-0R0	699	.2629	.0934	.0192	.0141
6	W-22-00-000	696	.1392	.2291	.0139	.0291
7	B-22-00-000	534	.2040	.0893	.0185	.0194
8	W-**-00-0R0	483	.1784	.1361	.0194	.0259
9	B-11-00-V00	445	.2327	.0695	.0222	.0186
10	W-33-00-*00	428	.0355	.0540	.0077	.0569
11	B-22-00-V00	398	.1230	.1040	.0172	.0324
12	W-11-00-00D	397	.2314	.1418	.0226	.0231
13	B-22-00-0R0	362	.1593	.0472	.0210	.0249
14	W-34-00-*00	315	.0345	.2159	.0102	.0938
15	W-22-00-00D	312	.2363	.2184	.0284	.0324
16	B-**-10-000	286	.3375	.0758	.0324	.0165
17	W-**-10-000	291	.2584	.1161	.0281	.0219
18	B-**-1I-0R0	259	.2553	.0532	.0333	.0204
19	W-**-10-V00	261	.1186	.0970	.0183	.0248
20	B-**-10-V*0	254	.1846	.0729	.0258	.0272
21	B-12-3I-V*0	228	.1952	.0182	.0177	.0213
22	B-*2-1I-V*0	212	.2305	.0387	.0325	.0231
23	W-22-00-V0D	230	.0840	.0370	.0161	.0444
24	B-**-10-0R0	201	.2763	.0770	.0359	.0224
25	W-22-00-0R0	203	.1229	.2251	.0263	.0599
26	B-**-1I-000	175	.3709	.0961	.0475	.0219
27	B-12-3I-000	164	.4043	.0921	.0511	.0220
28	W-11-00-0RD	173	.1935	.2253	.0363	.0497
29	B-**-00-00D	169	.2372	.1049	.0362	.0320
30	W-12-3I-0*0	160	.2260	.0300	.0360	.0246
31	W-*3-**-V*0	167	.0819	.0528	.0193	.0547
32	W-**-10-00D	156	.2845	.1239	.0403	.0293
33	B-**-*0-0*D	151	.2646	.0732	.0399	.0270
34	B-33-00-*00	162	.0664	.1094	.0184	.0712
35	B-12-3I-010	146	.1825	.0033	.0352	.0300
36	W-11-00-V0D	151	.1863	.1875	.0340	.0480
37	W-12-3I-V*0	142	.2192	.0771	.0374	.0323
38	B-34-00-V00	136	.0487	.5544	.0245	.2219
39	W-**-10-**D	128	.2528	.1142	.0412	.0357
40	B-12-3I-**D	120	.3837	.1442	.0634	.0314
41	B-22-3I-**0	116	.3427	.1103	.0614	.0298
42	B-*3-3I-V*0	106	.1113	.0394	.0298	.0560
43	W-**-1I-*00	105	.2739	.1149	.0501	.0348
44	W-**-10-0R0	80	.2377	.1458	.0598	.0495
Average For All Groups			.2066	.1163	.0290	.0371

Appendix F: CLUSTERS OF GROUPS - MLE's of μ and δ

** WHITES, 17-29 **

A. WHITE, FIRST ARREST, 17-19

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
1	2700	W-11-00-000	.2583	.1909	.0094	.0099
12	397	W-11-00-00D	.2314	.1418	.0226	.0231
8	464	W-**-00-0R0	.1784	.1361	.0194	.0259
28	173	W-11-00-0RD	.1935	.2253	.0363	.0497
3	888	W-11-00-V00	.1763	.1654	.0135	.0209
36	151	W-11-00-V0D	.1863	.1875	.0340	.0480
TOT	4773		.2279	.1804	.0068	.0078

B. WHITE, FIRST ARREST, 20-29, NOT VIOLENT

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
6	696	W-22-00-000	.1392	.2291	.0139	.0291
15	303	W-22-00-00D	.2363	.2184	.0284	.0324
25	203	W-22-00-0R0	.1229	.2251	.0263	.0599
TOT	1202		.1604	.2312	.0116	.0207

C. WHITE, FIRST ARREST, 20-29, VIOLENT

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
4	790	W-22-00-V00	.0757	.1061	.0087	.0309
23	230	W-22-00-V0D	.0840	.0370	.0161	.0444
TOT	1020		.0770	.0849	.0076	.0255

D. WHITE, FEW PRIOR ARRESTS, 17-29

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
17	291	W-**-10-000	.2584	.1161	.0281	.0219
32	156	W-**-10-00D	.2845	.1239	.0403	.0293
44	809	W-**-10-0R0	.2377	.1458	.0598	.0495
39	2619	W-**-10-**D	.2528	.1142	.0412	.0357
19	1054	W-**-10-V00	.1186	.0970	.0183	.0348
43	1283	W-**-1I-*00	.2739	.1149	.0501	.0348
TOT	1021		.2198	.1216	.0138	.0132

E. WHITE, MANY PRIOR ARRESTS, 17-29

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
30	160	W-12-3I-0*0	.2260	.0300	.0360	.0246
37	142	W-12-3I-V*0	.2192	.0771	.0374	.0323
TOT	302		.2211	.0489	.0249	.0174

**** BLACK & WHITE, 30+ ****

F. BLACK & WHITE, FIRST ARREST, 30-39

FREQ RANK	N	ATTRIBUTE ID	MLE μ	MLE δ	ST. DEV μ	ST. DEV δ
10	428	W-33-00-*00	.0355	.0540	.0077	.0569
34	162	B-33-00-*00	.0664	.1094	.0184	.0712
TOT	590		.0433	.0734	.0074	.0449

G. BLACK & WHITE, FIRST ARRESTS, 40+

FREQ RANK	N	ATTRIBUTE ID	MLE μ	MLE δ	ST. DEV μ	ST. DEV δ
14	315	W-34-00-*00	.0345	.2159	.0102	.0938
38	136	B-34-00-V00	.0487	.5544	.0245	.2219
TOT	451		.0370	.2846	.0094	.0859

H. BLACK & WHITE, MANY PRIOR ARRESTS, 30-39

FREQ RANK	N	ATTRIBUTE ID	MLE μ	MLE δ	ST. DEV μ	ST. DEV δ
31	167	W-*3-**-V*0	.0819	.0528	.0193	.0547
42	106	B-*3-3I-V*0	.1113	.0394	.0298	.0560
TOT	273		.0900	.0500	no convergence[44]	

[44]The maximum likelihood procedure did not converge in this case and so the standard deviation was not obtained. The μ and δ reported had the highest likelihood found by the algorithm, however, without convergence these estimates are not necessarily the maximum.

** BLACK, 17-29 **

I. BLACK, 17-29, FIRST ARREST

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
2	1151	B-11-00-000	.3128	.0479	.0146	.0081
29	169	B-**-00-00D	.2372	.1049	.0362	.0320
5	699	B-11-00-0R0	.2629	.0934	.0192	.0141
33	151	B-**-*0-0*D	.264[illegible]	.0732	.0399	.0270
9	445	B-11-00-V00	.2327	.0695	.0222	.0186
7	534	B-22-00-000	.2040	.0893	.0185	.0194
13	362	B-22-00-0R0	.1593	.0472	.0210	.0249
11	399	B-22-00-V00	.1230	.1040	.0172	.0324
TOT	3910		.2432	.0804	.0075	.0061

J. BLACK, 17-29, FEW PRIOR ARRESTS

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
16	286	B-**-10-000	.3375	.0758	.0324	.0165
24	201	B-**-10-0R0	.2763	.0770	.0359	.0224
20	254	B-**-10-V*0	.1846	.0729	.0258	.0272
26	175	B-**-1I-000	.3709	.0961	.0475	.0219
18	259	B-**-1I-0R0	.2553	.0532	.0333	.0204
22	212	B-*2-1I-V*0	.2305	.0387	.0325	.0231
TOT	1387		.2717	.0721	.0139	.0088

K. BLACK, 17-29, MANY PRIOR ARRESTS

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
27	164	B-12-3I-000	.4043	.0921	.0511	.0220
40	120	B-12-3I-**D	.3837	.1442	.0634	.0314
35	146	B-12-3I-0R0	.1825	.0033	.0352	.0300
21	228	B-12-3I-V*0	.1952	.0182	.0177	.0213
41	116	B-22-3I-**0	.3427	.1103	.0614	.0298
TOT	774		.2750	.0690	.0190	.0046

Q. BLACK, 17-29 (I,J, and K COMBINED)

FREQ			MLE		ST. DEV	
RANK	N	ATTRIBUTE ID	μ	δ	μ	δ
	6071		.2523	.0761	.0062	.0046

Appendix G: Regression Study of Variation in $\hat{\delta}$ Across Attributes

This appendix presents the procedure used to identify the following summary regression model for $\hat{\delta}$:

(G-1)

$$\hat{\delta} = \underset{(.01)}{.18} - \underset{(.01)}{.11}\, X_b + \underset{(.11)}{.09}\, X_1 - \underset{(.02)}{.06}\, X_{wf} - \underset{(.03)}{.13}\, X_{wm} - \underset{(.03)}{.10}\, X_{wtv} + \epsilon$$

N = 44

SSE = 61.4

R^2 = .7650

Where,

$$X_b = \begin{cases} 1 & \text{black and age} < 40 \\ 0 & \text{else} \end{cases}$$

$$X_1 = \begin{cases} 1 & \text{aged 40+} \\ 0 & \text{else} \end{cases}$$

$$X_{wf} = \begin{cases} 1 & \text{white, few prior arrests, age} < 40 \\ 0 & \text{else} \end{cases}$$

$$X_{wm} = \begin{cases} 1 & \text{white, many prior arrests, age} < 40 \\ 0 & \text{else} \end{cases}$$

$$X_{wtv} = \begin{cases} 1 & \text{white, no prior arrests, violent, aged 20-39} \\ 0 & \text{else} \end{cases}$$

Procedure:

Several series of regression models are considered to explain the variation in $\hat{\delta}$ as a linear combination of the attributes and their interaction terms. To account for heteroskedasticity, weighted least squares is used. The weight associated with a group is the estimated precision of the maximum likelihood estimate where precision is defined as the inverse of the estimated standard error squared.

First, models which include no interaction terms among the attributes are considered in order to identify those attributes with significant main effects on $\hat{\delta}$. Then, models which consider interactions between the primary effects identified in the first models and the other attributes are considered. Finally, the thoroughness of the chosen model is confirmed using plots of $\hat{\delta}$ versus $\hat{\mu}$ to examine the range of variation of $\hat{\delta}$ within the identified clusters.

When possible, models are compared using the standard F-statistic (Equation G-2). This statistic tests whether the difference in the explanatory value of 2 models is significant. This statistic is appropriate for pairs of models that share the same set of parameters where in one model some of the parameter coefficients are restricted to be zero.

(G-2)

$$F = \left(\frac{(SSE_r - SSE_c)/(k-g)}{(SSE_c)/(n-k)} \right) \sim f_{k-g,n-k}$$

Where,

SSE_r = sum of squared errors restricted model

SSE_C = sum of squared errors unrestricted model

g = # of parameters in restricted model

k = # of parameters in complete model

n = # of groups

If $F > f_{k-g,n-k}(\alpha)$ then it can be concluded at the α level of significance that the unrestricted model has better explanatory value than the restricted model. Conversely, if $F < f_{k-g,n-k}(\alpha)$ then the hypothesis that the restricted model provides less explanatory value can not be rejected at the α level of significance.

Main Effects Models:

The first model explores the main effects on $\hat{\delta}$ of offender attributes. The intercept of this model is the termination rate of young, white, first-time, property offenders (W-11-00-000). To identify main effects, twelve binary variables, one for each additional level of the 8 attributes, are included as independent variables. Table G-1 reports that the only main effect associated with significant observed variation in $\hat{\delta}$ is RACE. Black offender groups exhibited a lower termination rates overall than white offender groups.

Table G-1: All Main Effects Regression Model

Indep. Var.	Coefficient Est.	(S.E.)	t-test
INTERCEPT	.157	(.013)	12.21*
RACE=B	-.080	(.014)	-5.87*
AGE1=2	.048	(.036)	1.34
AGE1=3	-.001	(.113)	- .01
AGENOW=2	-.036	(.034)	-1.05
AGENOW=3	-.049	(.080)	- .61
AGENOW=4	.126	(.191)	.66
CPRIOR=1	-.015	(.018)	- .83
CPRIOR=3	-.015	(.040)	- .37
IPRIOR=I	.014	(.027)	.50
VEVER=V	-.015	(.018)	- .84
REVER=R	.004	(.016)	.26
DEVER=D	.014	(.020)	.67

*significant at the α=.05 level

N = 44
SSE = 98.0
R^2 = .6248

The parameter associated with offender groups aged 40+ years is quite large (.13), though, this result is not statistically significant at an α=.05 level due to the large standard errors associated with $\hat{\delta}$ for the two groups of 40+ year-old offenders:

Group	$\hat{\delta}$	S.E.
W-34-00-*00	.22	(.09)
B-34-00-V00	.55	(.22)

However, since $\hat{\delta}$ is considerably larger for these two groups, they are classified as a separate cluster, even though the main effects parameter associated with older (40+) offenders is not significant. Further regression analyses reported in this appendix exclude these two groups.

The results of this first main effects model indicates that a RACE-only model should be considered. Table G-2 reports that the regression with only one explanatory variable, RACE, accounts for 51% of the observed variation.

Table G-2: RACE-Only Regression Model

Attribute Level	Est.	(S.E.)	t-test
INTERCEPT	.150	(.010)	14.83*
RACE=B	-.083	(.013)	-6.56*

*significant at the α=.05 level

N = 42
SSE = 123.1
R^2 = .5179

An F-test shows that given the reduction in the number of degrees of freedom, the RACE-only model does not perform significantly worse than the complete main effects model at the

α=.05 level[45]:

$$F = \frac{(128.5 - 98.0)/(13 - 2)}{98.0 \quad /(44 - 13)} = \frac{2.77}{3.16} = .87$$

$$f_{11,31}(.05) = 2.05$$

$$F < f_{11,31}$$

To test if a regression model with two main effects could improve upon the RACE-only model a series of regression analyses is performed in which RACE and each of the 7 other primary attributes are used as explanatory variables for $\hat{\delta}$, one attribute at a time. A comparison of the RACE-only model with these two main effects models indicates that additional main effects do not significantly help to explain the variation in $\hat{\delta}$ (see Table G-3):

[45]For the purpose of comparison with the all main effects model the results of the RACE-only model using all 44 groups which has a SSE of 128.5. The results reported in Table G-2 are for the 42 groups excluding groups of offenders aged 40+.

Table G-3: Two Main Effects Regression Models Which Include RACE

Attribute	R^2	SSE	DF	F-Value[46]
AGE1	.5364	118.4	2,38	.77
AGENOW	.5439	116.5	2,38	1.09
CPRIOR	.5758	108.4	2,38	2.59
IPRIOR	.5497	115.0	1,39	2.78
VEVER	.5484	115.3	1,39	2.67
REVER	.5182	123.1	1,39	.03
DEVER	.5232	121.8	1,39	.45

The failure of these additional attributes to significantly improve the RACE-only model indicates that if the attributes have further explanatory power it may differ across black and white offender groups.

Interaction Models:

For further analyses, the data are divided into two groups, black and white offenders. For black offender groups, none of the 7 remaining attributes has significant explanatory value. Table G-4 reports the results of F-tests for a series of regressions in which each attribute is used one at a time as an explanatory variable of $\hat{\delta}$ for black, 17-39 year-old offender groups.

[46]This F-test compares the SSE for the RACE-only model presented in Table G-2 to the RACE and a second attribute regression model for each of the attributes. A two main effects regression model with significantly more explanatory value than the RACE-only model at the α=.05 level would have an F-Value greater than 3.25 for tests with (2,38) degrees of freedom and 4.10 for (1,39) degrees of freedom. AGENOW provide only 2 additional degrees of freedom since groups of offenders aged 40 years and older (AGENOW=4) are excluded based on the main effects analysis.

Table G-4: Summary of Single Effect Regression Models for Black Offender Groups Aged 17-39

Attribute	R^2	**SSE**	**DF**	**F-Value**[47]
AGE1	.0704	35.2	2,18	.68
AGENOW	.0004	37.8	2,18	.00
CPRIOR	.0036	37.7	2,18	.03
IPRIOR	.0025	37.8	1,19	.05
VEVER	.0465	36.1	1,19	.35
REVER	.0020	37.8	1,19	.04
DEVER	.1298	32.9	1,19	2.83

Further confirmation of the similarity of termination rates among groups of black, 17-39 year-old offenders is evident in a plot of $\hat{\mu}$ vs. $\hat{\delta}$. Figure G-1 reports that the termination rates for black offender groups are primarily between .07 and .11 per year. Based on the single attribute regression models and the confirmatory plot of $\hat{\mu}$ vs. $\hat{\delta}$, black 17-39 year-old offender groups are included in a single cluster.

***** Figure G-1 here *****

Figure G-1: $\hat{\mu}$ vs. $\hat{\delta}$ for Black 17-39 Year-old Offender Groups

[47] The F-test reported compares the SSE for the single attribute model to an intercept-only model which has SSE = 37.9. An attribute which significantly improve the model at the α=.05 level would have an F-value greater than 3.49 for tests with (2,18) degrees of freedom and 4.35 for (1,19) degrees of freedom.

A similar series of single effect regression models for white, 17-39 year-old, offender groups (see Table G-5) indicates that the termination rate decreases as the number of prior arrests increases (CPRIOR).

Table G-5: Summary of Single Effect Regression Models for White Offender Groups Aged 17-39

Attribute	R^2	SSE	DF	F-Value[48]
AGE1	.0513	80.9	2,18	.62
AGENOW	.1916	68.9	2,18	2.13
CPRIOR	.5850	35.3	2,18	12.68*
IPRIOR	.3417	56.1	1,19	9.86*
VEVER	.0923	77.4	1,19	1.93
REVER	.0000	85.3	1,19	.00
DEVER	.0010	85.2	1,19	.02

*significant at the α=.05 level

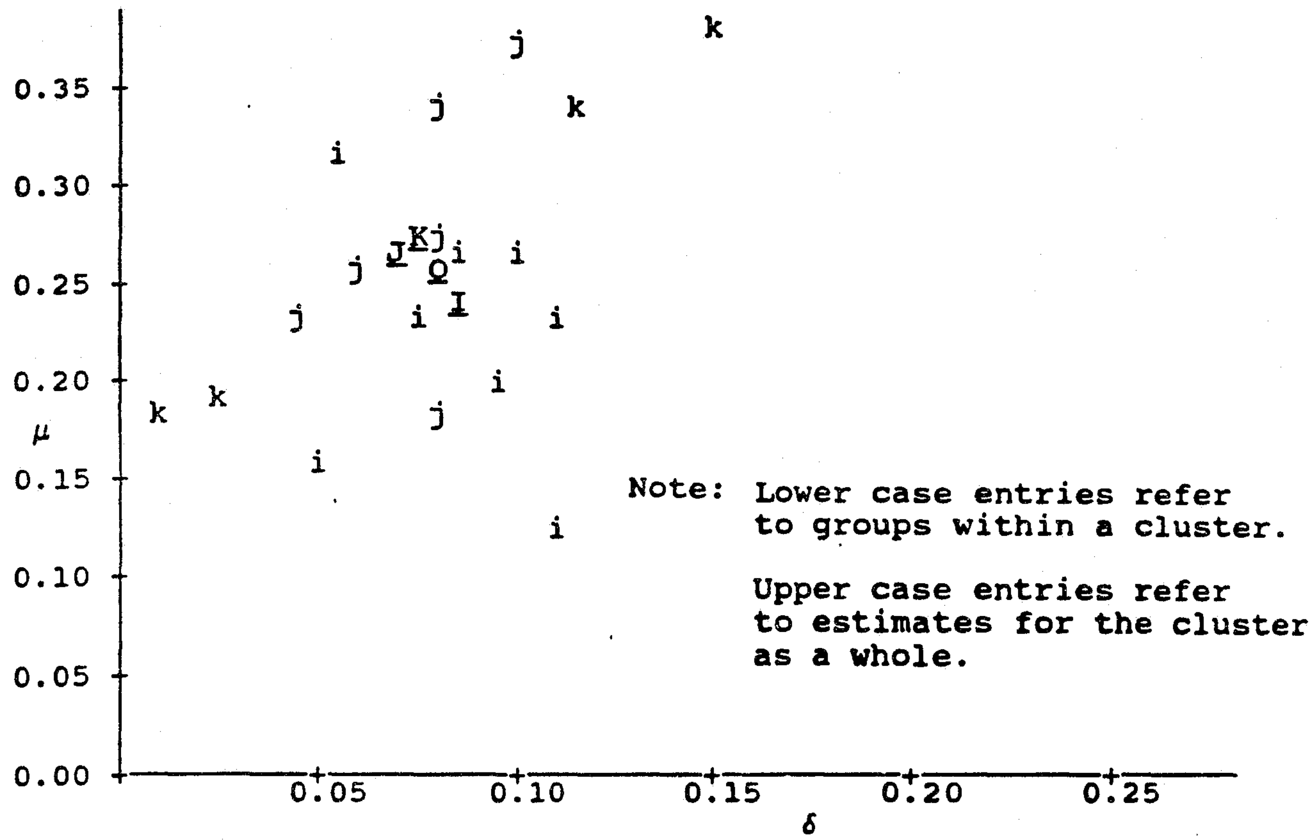
The improved explanatory power associated with both CPRIOR and IPRIOR are significant at the α=.05 level. However, IPRIOR is highly correlated with CPRIOR -- only offenders with a prior arrest can have a record of prior incarceration -- and so it suffices to consider only CPRIOR at this point. Figure G-2 reports that the termination rate decreases as the number of prior arrests increases with the exception of 20-39 year-old, first-time, violent

[48]Similar to the test in Table G-4, the F-test for this model compares the SSE for the single attribute model versus an intercept only model with SSE = 85.3. An attribute which significantly improves the model at the α=.05 level would have an F-Value greater than 3.49 for tests with (2,18) degrees of freedom and 4.35 for (1,19) degrees of freedom.

offenders.

Figure G-2: $\hat{\mu}$ vs. $\hat{\delta}$ for White 17-39 Year-old Offender Groups **** place here ****

Further analysis of G-2 reveals that little variation in $\hat{\delta}$ exists within individual prior groups with the single exception previously noted. Based on this analysis white 17-39 year-old offenders are divided into 4 clusters: 20-39 year-old violent first-time offenders; other first time offenders; offenders with few prior arrests; and offenders with many prior arrests.

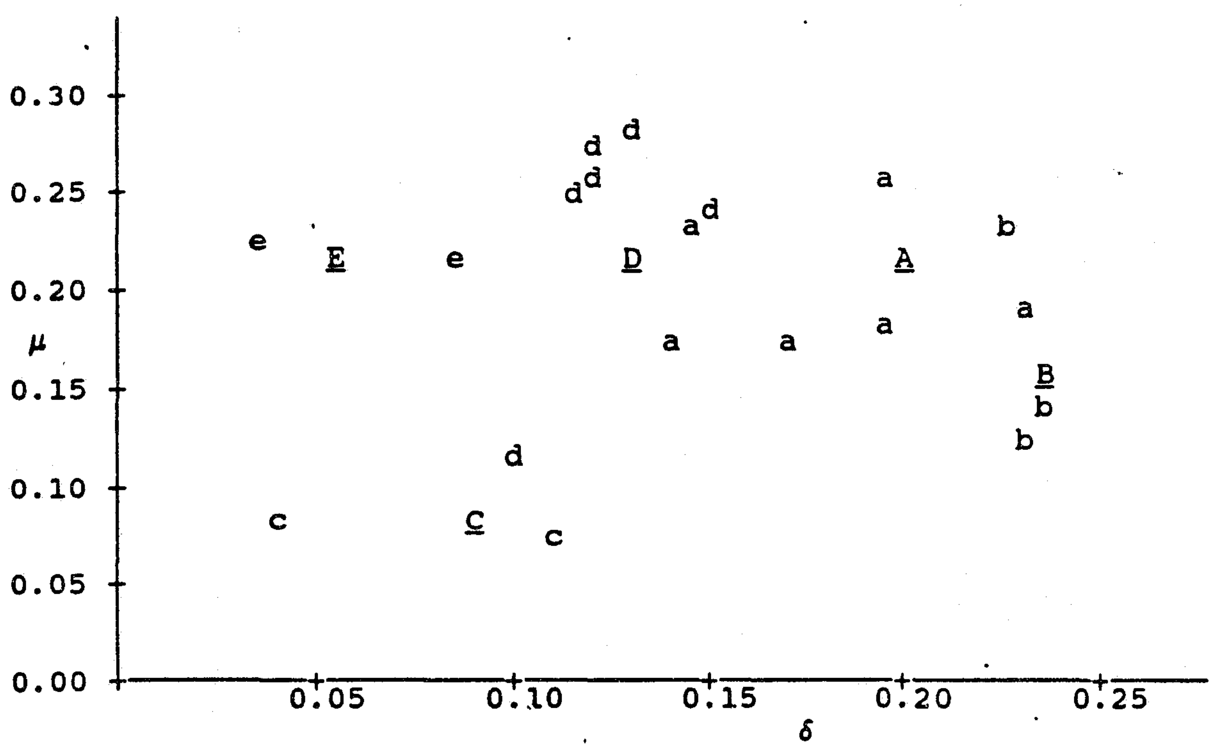


	CLUSTER	δ	μ
I.	BLACK, 17-29, FIRST ARREST	.08	.24
J.	BLACK, 17-29, 1-2 PRIORS	.07	.27
K.	BLACK, 17-29, 3+ PRIORS	.07	.28
Q.	BLACK, 17-29 (I,J,K COMBINED)	.08	.26

Figure G-1: $\hat{\mu}$ vs. $\hat{\delta}$ for Black 17-39 Year-Old Offender Groups

Note: Lower case entries refer to groups within a cluster.

Upper case entries refer to estimates for the cluster as a whole.



	CLUSTER	δ	μ
A.	WHITE, FIRST ARREST, 17-19	.18	.23
B.	WHITE, FIRST ARREST, 20-29, NON-VIOLENT	.23	.16
C.	WHITE, FIRST ARREST, 20-29, VIOLENT	.08	.08
D.	WHITE, 1-2 PRIORS, 17-29	.12	.22
E.	WHITE, 3+ PRIORS, 17-29	.04	.22

Figure G-2: $\hat{\mu}$ vs. $\hat{\delta}$ for White 17-39 Year-Old Offender Groups