152464

AT RISK OF REARREST FOR A VIOLENT CRIME-- PREDICTING HIGH-STAKES, HIGH-SPEED RECIDIVISM: DEVELOPING PREDICTION MODELS IN TWO BIRTH COHORTS

Neil Alan Weiner

Senior Research Associate
Center for the Interdisciplinary Study of Criminal Violence
Sellin Center for Studies in Criminology and Criminal Law
The Wharton School
University of Pennsylvania
436 Vance Hall
Philadelphia, PA 19104-6301

(215) 898-6590

This final report was prepared under grant #86-IJ-CX-0052 from the National Institute of Justice, Offender Classification and Prediction of Criminal Behavior Program, U.S. Department of Justice. Points of view or opinions presented herein are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Figures

Chapter 1

AT RISK OF REARREST FOR A VIOLENT CRIME--
PREDICTING HIGH-STAKES, HIGH-SPEED RECIDIVISM:
HOW WELL CAN WE DO IT?

THE STUDY GOALS IN A CAPSULE

This study investigates some basic aspects of our present capacity to predict individual and aggregate arrests for serious violent crimes. Two notions are central to this investigation: the probability and the timing of arrests. The study first looks at how well we can predict whether and when an individual who has been arrested for a serious violent crime will be arrested again for one of these same crimes. The study then looks at how well we can predict the aggregate number of individuals who, having been arrested for serious violent crimes, will be arrested again for these same kinds of crimes by some time point of interest.

The main study goal is to improve individual- and aggregate-level prediction within a public policy framework. We want to see whether it is possible to strengthen the decision making capabilities of those justice system officials who must make front line decisions about how to deal with persons who have just been arrested for serious violent crimes.

Judging from past experience in criminology and criminal justice, and in the social sciences more broadly, we are unlikely to achieve a spectacularly high level of predictive accuracy. Moderate, not quantum, advancements are the rule. And, even moderate advancement tends to be hard won and subject to later erosion. Predictive accuracy, whatever level is achieved, can be expected to recede because of natural changes in the very phenomena which one tries to predict; basically, at some point in time, the prediction tool no longer fits, or matches, the evolving predicted phenomenon. The natural slide toward predictive obsolescence argues, however, for periodically updating one's prediction tool, not for scraping that tool.

The statistical prediction of arrests for violent crimes is now at a virtual standstill. This study tries to identify and overcome some of the technical obstacles to predicting whether and when these arrests will occur, to undo some of the research inertia in this area, and to harness findings in an applied setting.

SETTING THE STAGE: RESPONDING TO A NATIONAL PROBLEM OF MAJOR PROPORTIONS

The problem of violent crime in the United States is grave, has been for some time, and is unlikely to abate soon. The widely used and useful yardstick of official crime statistics, the Federal Bureau of Investigations's Uniform Crime Reports (UCR), tells a sobering story of criminal violence in the nation. The most recent UCR estimates, for 1989, indicates that the police nationwide edged close to making 700,000 arrests for serious violent index crimes--more than 22,000 for murders, nearly 40,000 for forcible rapes, more than 165,000 for robberies, and almost 460,000 for aggravated assaults.[1]

---

[1] The most recent national estimates, for 1989, can be found in the U. S. Department of Justice, Federal Bureau of Investigation. 1990. Uniform Crime Reports for the United States, 1989. Washington, DC: U. S. Government Printing Office, table 24, p. 172. These totals were based on all reporting agencies and on estimates for nonreporting agencies. Excluding the estimates calculated for the nonreporting agencies, there were approximately 537,000 reported arrests for violent index crimes: about 18,000 for murder and nonegligent manslaughter, 30,000 for forcible rapes, 134,000 for robbery, and 355,000 for aggravated assault. Males accounted for about 90 percent of the arrests; juveniles and young adults, ages 10 to 25, accounted for about 50 percent of the arrests; and whites accounted for about 51 percent, blacks for more than 45 percent, and other racial and ethnic minorities for less than 2 percent. See U. S. Department of Justice, Federal Bureau of Investigation. 1990. Uniform Crime Reports for the United States, 1989. Washington, DC: U. S. Government Printing Office, tables 33, 36-38, pp. 182-83, 188-92.

For each of these crimes, arrest rates were at their peak levels in our nation's most populated cities.[2] And, for more than two decades now, these rates have been highest among: males; nonwhites, mainly blacks; and during late adolescence and early adulthood, stretching from ages 16 to 25.[3] These official statistics lay the bulk of the problem of serious violent crime squarely on the doorstep of young minority males living in our nation's urban centers.

These urban hot spots are among the most financially beleagured in the nation. Public resources are now scarce in these areas, and they are likely to become increasingly so in the next decade. As the competition for these scarce resources heats up among governmental agencies and their constituencies, the capacity of the nation's two tiered justice system (the js), the juvenile justice system (the jjs) and the adult criminal justice system (the cjs), to staunch the current crest of criminal violence will be weakened. Js personnel will increasingly have to make hard policy choices, even harder than the ones they now have to make, about how to dispose of violent criminals. To the extent that the nation drifts in a more punitive direction, these hard choices will also become increasingly harsh in their impact.

But what decision choices should these js gatekeepers--police, prosecutors and lower court judges--make in response to these violent persons

---

[2] U. S. Department of Justice, Federal Bureau of Investigation. 1990. Uniform Crime Reports for the United States, 1989. Washington, DC: U. S. Government Printing Office, table 26, p. 174.

[3] U. S. Department of Justice, Federal Bureau of Investigation. 1990. Age-Specific Arrest Rates and Race-Specific Arrest Rates for Selected Offenses, 1965-1988. Washington, DC: U. S. Government Printing Office, pp. 17-84, 298-300, 332-36.

when they are arrested? Certainly one of the main threshold decisions which must be made, even when that decision seems a foregone conclusion because of the uniform gravity of violent index crimes, is whether to confine the arrested person in a secure facility prior to trial or whether to release the person on bail.[4] But this decision making is neither simple nor straightforward. Should the preferred decision choice change as the person accumulates more arrests? Should the fact that a weapon was used--for example, a firearm as opposed to a knife--influence the decision maker? What should the decision maker do if the arrested person began notching arrests very early in life? And what should the decision maker do with information about extra-legal characteristics of the arrested person (e.g., race, socioeconomic status). These kinds of questions can easily multiply, their number and nature depending upon several things: the substance and perceived soundness of the theories of violent behavior which are embraced by the decision maker, the major research findings on violent behavior to which the the decision maker has been exposure, and professional experience and wisdom. Clearly, the decision choices are complex because of the quite different types, causes, and circumstances of the violent criminal behavior which must be taken into account. Just as clearly, the decision choices are uncertain because of our incomplete knowledge about the causes and courses of nearly every type of criminal violence.

---

[4] Another critical js decision relates to the selection of offenders for priority prosecution. See Marcia Chaiken and Jan Chaiken, Redefining the Career Criminal: Priority Prosecution of High-Rate Dangerous Offenders (Washington, DC: U.S. Department of Justice, National Institute of Justice, 1990). While the specific policy and administrative issues may be different when studying the decision to impose detention as opposed to selecting a criminal for priority prosecution, both kinds of studies use prediction-based classification to bolster the accuracy of decision making.

One critical component of the threshold decision either to confine or to release an arrested person relates to the risk which officials believe the person runs of being rearrested for a serious violent crime; a person at high risk of rearrest is, more often than not, probably at high risk of confinement, all other things being equal (e.g., media publicity). The decision, however, to confine someone in a secure facility in order to avert the commission of serious violent crimes when it is unlikely that these crimes will be committed can sting in two ways: it impinges unnecessarily upon that individual's liberty, and it wastes precious and scarce public resources. It is certainly possible, however, to reduce both the impingement of liberty and the wastage of resources by developing explicit and formal statistical tools which improve the accuracy with which we can predict whether a person will be rearrested for a serious violent crime. The present study investigates whether this type and application of prediction might be made accurate enough to be of practical value to js officials as they routinely discharge their decision making responsibilities.

The decision making quandary faced by js officials is not unlike that faced by a savvy card player who wagers a bet on a particular card hand: the card player must decide whether there exists a sufficient amount of information to predict with substantial accuracy the outcome of that card hand? If the card player is convinced that this information exists, is presently available, and ensures sufficiently high predictive accuracy, the wager is (usually) on. The merits of the particular card hand being played, however, is commonly weighed by the more adept players within the context of long term success. While the card player seeks to maximize the chances of winning each hand, no card player expects actually to do so. The best players understand that individual card hands which are lost, or even a run of lost hands, may be the unavoidable but necessary links in a chain of play that is an overall success. Indeed, the lost hand or run of bad luck often provides new information that can sharpen future play. The astute player finds ways of profiting from mistakes.

This study examines whether certain kinds of information commonly available to front line js decision makers yield sufficiently high predictive accuracy with respect to rearrests for serious violent crimes to support the js wager that confining an offender in a secure facility will likely payoff in averting at least one more violent crime. And this wager must be viewed from a long term perspective; lost wagers are sometimes to be expected (and absorbed), even when wagering is likely to be successful over the long run. In view of the potentially steep stakes involved in this wager--foregone public protection or inpingement of individual liberty--the decision maker must have enough relevant predictive information at hand to convince himself that it is worthwhile to shift from risk aversion to risk taking.

Cjs decision making is, then, a risky business. That is especially so with respect to violent criminals. Officials can be and, indeed, often are badly burned by some of their predictions--even when they have played their cards exactly right. This study shows that this type of outcome is inescapable, that certain types of prediction can nevertheless still be quite useful, and that even inaccurate predictions yield information which is potentially useful to improving prediction accuracy.

THE TYPES, USES, ERRORS, AND COSTS OF VIOLENCE PREDICTION: INDIVIDUAL DECISIONS AND INSTITUTIONAL POLICY DECISIONS

Types of Prediction

Just how accurately can officials at the front lines of our nation's js identify which persons who have just been arrested for serious violent crimes will be rearrested for one of these same crimes? This question broadly expresses a classic problem in prediction-based classification: first, calculating the numerical risk that a person will engage in a particular behavior or experience a particular event or condition (e.g., the risk of rearrest for a violent crime); second, assigning that risk to one of two (or more) discrete, mutually exclusive outcome categories which numerically classify the level of risk (e.g., high risk of rearrest for a violent crime versus low risk of rearrest for a violent crime); and, third, deciding whether the level of assigned risk constitutes sufficient grounds for predicting that the particular outcome will occur (e.g., high risk of rearrest for a violent crime leads to the decision to predict that a rearrest will occur, whereas a low risk of rearrest for a violent crime leads to the opposite decision).

To illustrate some key features of prediction-based classification, consider the following common example relating to the administration of the jjs: One goal of detaining a delinquent in a secure facility who has been apprehended for involvement in serious violent behavior is to avert further involvement in such serious behavior while the youth is awaiting the adjudication trial and disposition. To make a prudent decision about whether to detain the delinquent, a juvenile court judge might want explicitly and formally to assess the behavioral risk that that delinquent either will be rearrested for a serious violent crime or will not be rearrested for a serious violent crime. In making this assessment, the judge marshals and weighs various pieces of information assumed or demonstrated to be related to the risk of the behavioral outcome, usually according to some quantitative procedure which can be used to classify the level of risk. If the calculated risk, measured as a probability, exceeds a prespecified classification cutpoint (the "cutting score"), let us say, a .90 probability, then the judge predicts that the delinquent will be rearrested; if the risk falls at or below the classification cutpoint, then the judge predicts otherwise. In the above example, if the calculated behavioral risk indicates rearrest for a serious violent crime, the judge might decide to detain the delinquent in secure custody. From an administrative standpoint, the judge has employed prediction-based classification to screen a delinquent who has just been arrested for a serious violent crime in order to decide whether that delinquent qualifies for a particular disposition: predicting that a particular outcome will occur triggers the decision to choose a particular disposition, presumably in order to secure a rational organizational objective, in the example, averting future involvements in violent crimes.

Risk assessment (calculation of the risk level), risk assignment (classification of the risk level), risk prediction (decision choice about the likely behavioral outcome based on the risk assignment), and prediction evaluation (accuracy of the prediction) are prominent aspects--the first three anyway--of explicit and formal decision making in organizations like the jjs and the cjs which process persons through sequential stages, each of which branches into discrete, mutually exclusive dispositional pathways.[5] But risk

[5] Prediction-based classification involves five steps: (1) specifying a set of two (or more) mutually exclusive behavioral outcomes (e.g., rearrest for a violent crime or no rearrest for a violent crime), (2) stipulating a

assessment and its collateral procedures have certain conceptual facets which may not be apparent at first glance, and these facets may be quite important both to the ultimate technical precision and practical accuracy of the risk assessment, and to the organizational application of that assessment. Some risk assessment techniques may, for instance, utilize more of the available information about the behavioral outcome, buttressing the reliability of results, and permit more versatile organizational applications, stimulating and underwriting their practical adoption.

It may not have been apparent, but the prediction example presented above, relating to the juvenile court judge's decision about whether to detain a delinquent, actually involved the notion of time. If a seriously violent delinquent is likely to be arrested once again for a serious violent crime, when will that be? Unfortunately, classical prediction-based classification techniques fail formally to take time into account, information which can almost always be retrieved from official records. This technical gap can severely limit the utility of these techniques for both theory and public-policy development. However, by formally taking time into account, one can refine the earlier prediction-based classification question, with some useful and challenging results: With what degree of accuracy can our nation's front line, js officials identify those persons who have just been arrested for

---

probability cutpoint which divides the behavioral outcomes into discrete categories (e.g., a probability greater than .90 results in predicting a rearrest for a violent crime, whereas a probability less than or equal to .90 results in the opposite prediction), (3) developing a procedure, usually expressed as a statistical equation, to calculate the probability of a behavioral outcome, (4) assigning a person to one of the outcomes depending upon whether the calculated probability falls above or below the probability cutpoint, and (5) evaluating whether the predicted outcome corresponds to the actual outcome (e.g., if a person is predicted to be rearrested for a violent crime, was that person rearrested).

serious violent crimes who will be rearrested for one of these same crimes within a specified amount of time? This question expresses a much more contemporary prediction-based classification problem, one which has undoubtedly, if often only tacitly, for some time been on the minds of both js researchers and practitioners: the assessment of individual risk involving a continuous behavioral outcome, in the present study, the time until rearrest for a violent crime.

Recall the earlier example of the juvenile court judge: The judge might want explicitly and formally to assess the behavioral risk that the delinquent will be rearrested for another serious violent crime crime before a particular cutoff time, such as the scheduled date of the adjudication trial. If the calculated risk exceeds a stipulated classification cutpoint by that date, let's again say, a .90 probability of rearrest within three months, which is when the adjudication trial is scheduled, the judge predicts that the delinquent will be rearrested; as in the earlier example, if the risk falls at or below the classification cutpoint by the designated date, the judge decides otherwise. And, also as in that earlier example, each decision triggers the selection of a different disposition.

Prediction which evaluates the risk that a person will be rearrested for a serious violent crime by a particular time, time-related prediction, is potentially more useful to js practitioners and policy makers than prediction which ignores the time dimension, time-unrelated prediction. This is so because time-related prediction can yield richer, more versatile policy-relevant information than time-unrelated prediction: the risk of rearrest and the time period over which that risk level operates. It is encouraging that researchers have begun more widely to recognize the benefit of predicting both whether a person will be rearrested and, if that is likely to be, when that will be.[6]

The risk component of time-related prediction provides the technical means to sort violent criminals into groups having ascending probabilities of rearrest. As described in the above example, arrested persons located at the higher rungs of the risk gradient may require differential handling, perhaps more restrictive or harsh judicial dispositions (e.g., a longer placement in confinement) or the provision of more massive, diverse, and timely social services. If one eschews for ethical reasons using prediction-based classification to select arrested persons for the more restrictive or intensive dispositions, because these dispositions are considered unjustifiably punitive and impinging, then just the opposite decision choice might be made. Only those persons at the lower rungs of the risk gradient might be targeted for intervention, and they would receive the less restrictive and less intensive dispositions, for example, placement in nonsecured custodial care (i.e., selective deinstitutionalization).

The time component of time-related prediction provides the technical means to identify those time periods during which a person who has been arrested for a violent crime undergoes the more pronounced risks of rearrest. Resources can then be strategically applied during those time periods. However, the overall usefulness of a prediction scheme, such as one linking risk and time, can only be realized if the resultant predictions are sufficiently accurate.

---

[6] Blumstein, A., Cohen, J., Roth, J. A., Visher, C. A., eds. 1986. Criminal Careers and "Career Criminals". Washington, DC: National Academy Press; Schmidt, P., Witte A. D. 1988. Predicting Recidivism Using Survival Models. New York: Springer-Verlag.

The Uses of Prediction: Individual Decisions and Institutional Policy Decisions

Time-related predictions of rearrest can potentially pay off in two broad ways: decisions about individual dispositions and institutional policy decisions.[7] First, js officials can better protect the public from seriously violent criminals if these officials are able to predict more accurately which specific criminals arrested for violent crimes are at greatest risk of being rearrested for these same crimes within some specified time period, for instance, during the period right after arrest or between the time of arrest and the adjudication trial. The capacity to predict when the person is most likely to be rearrested can help js officials make more informed decisions about whether to detain criminals in secure facilities and, if detention is warranted, for how long? The police are the first to face the detention decision, and the lower court judges, at the first court appearance, are the second to face it. The public is protected to the extent that violent crimes which might otherwise be committed by this particular criminal are averted by the timely decision to detain that criminal. Increased public protection is one potential payoff of individual decision making.

Second, js administrators can better anticipate the overall institutional workload during a particular time period if they are able to forecast more accurately the aggregate number of violent criminals who will be rearrested for these crimes during that particular time period. More accurate

---

[7] Gottfredson, D. M. 1987. Prediction and classification in criminal justice decision making. In Crime and Justice, A Review of Research, vol. 9, Special Issue on Prediction and Classificiation--Criminal Justice Decision Making, ed. D. M. Gottfredson, M. Tonry, 1-20. Chicago, IL: University of Chicago Press; Gottfredson, M. R., Gottfredson, D. M. 1988. Decision Making in Criminal Justice: Toward the Rational Exercise of Discretion, 2nd Edition. New York: Plenum Press.

time-related predictions of aggregate rearrests can strengthen the credibility of strategic planning involving the allocation of js resources, such as the number of public prosecutors who might need to be assigned to the major crimes unit of the district attorney's office, and the level of government funds which might be needed to finance the construction of new secure jail cells to house violent criminals awaiting trial. More refined organizational planning is one potential payoff of institutional policy decisions.

The Errors and Costs of Prediction

Js practitioners can make two types of errors when predicting rearrests for violent crimes, and these errors can, in turn, have two important social and individual impacts: diminished law and/or diminished order. First, some violent criminals who are predicted not to be rearrested for violent crimes will in fact be rearrested for these crimes (the false negative predictions). Public safety (the order component) can diminish to the extent that these criminals are not confined in secure facilities or provided with intensive supervisory and supportive services which might avert the serious violent crimes which will result in their rearrest. Second, some violent criminals who are predicted to be rearrested for these crimes will in fact not be rearrested for them (the false positive predictions). Individual liberty (the law component) can diminish to the extent that these criminals are placed in secure facilities or involuntarily provided with intensive and intrusive supervisory or supportive social services when they would not have become reinvolved in serious violent crimes. Both public safety and individual liberty can erode to the extent that these two types of false predictions are made.

The problem of unduly restricted liberty would crop up, perhaps unavoidably so, even if reinvolvement in violent crime were capable of being predicted with virtual certainty. Friction would naturally arise between restricted liberty and the legal "presumption of innocence" which operates prior to both the occurrence of the predicted rearrest and the legal finding of guilt for that arrest. It may be, then, that whenever predictive decision making is employed by officials in our nation's js, social utility necessarily and quite rightly clashes with legal justice, resulting in their uneasy but, from the point of view of individual freedom, welcome and ultimately healthy tension.[8] Regardless of the direction of imbalance at any particular time, however, the tension is starkest and made more difficult ethically to defend when rooted in decision making error. Under any circumstances, it might be difficult for policy makers to embrace, for example, public safety over individual liberty, but that embracement may be rendered still more difficult knowing that it entails accepting the presence of false positives and the infringements of individual liberty following from these false predictions.

Because the social and individual stakes can be so high when the focus is serious violent crime, the tension between public safety and personal liberty is ratcheted upward yet another notch. Predicting that a person will be rearrested for a serious violent crime as opposed to a less serious crime

---

[8] The decision to impose pretrial detention, when based on a prediction-based classification scheme, will generally be legally sustained even if these decisions infringed on individual liberty as long as it can be shown that these decisions were not "intended to be punitive" but rather were incidental to some other legitimate purpose. For a detailed discussion of this and kindred legal and ethical issues in prediction and classification, see Tonry, M. 1987. Prediction and classification: legal and ethical issues. In Prediction and Classification: Criminal Justice Decision Making, Special Issue of Crime and Justice, A Review of Research, Vol. 9, ed. D. M. Gottfredson, M. Tonry, 367-413. Chicago, IL: University of Chicago Press.

is likely to place that person in greater jeopardy of loss of liberty, and for a longer period of time. The loss of liberty can result, in turn, in massive personal, socioeconomic, and legal liabilities (e.g., exposure to potentially adverse confinement conditions like stress and physical attacks; loss of job, income and educational opportunities; social ostracism and isolation; rupture of family and community ties; reduced capacity to participate in one's own defense). If the person will, in fact, not be rearrested for a serious violent crime, the false positive pitfall, that person may suffer greatly and uselessly, and, depending upon one's ethical position, also undeservedly. Similarly, predicting that a person will not be rearrested for a serious violent crime when that person will in fact be rearrested for such a crime, the false negative pitfall, heightens the danger that someone in the community will sustain grave physical injury or that the orderly pursuance of justice will be compromised (e.g., witness intimidation; failure to appear at the trial).

Aggregate prediction can also create tensions when js administrators formally integrate such prediction into their strategic planning. Overprediction can result in the unwarranted overdrawing of limited or scarce resources, perhaps impoverishing or dooming other needed and useful js programs and services. Underprediction can result in the unwarranted underdrawing of these resources, perhaps delaying the initiation of, or foreclosing on, the creation of beneficial js programs and services.

Prediction is clearly, then, a ubiquitous, central, challenging, and consequential aspect of js operations. Much of this prediction, however, is the "off-the-cuff" variety, based on unsystematic personal experience, institutional traditions, and sheer, because inexplicit, hunches. The present study takes the position that these kinds of subjective "intuitive" predictions and even the much more structured "clinical" predictions, while oftentimes quite functional and beneficial to both decision makers and their clients, are not as potent as objective, formal statistical approaches in reducing the uncertainty of the resultant predictions. We focus on the front-end decision points in the js, assessing in a <u>limited and preliminary</u> way whether optimism is warranted that pragmatically useful predictive accuracy can be achieved by formal statistical methods, especially by one of the more sophisticated recent approaches known as failure time analysis.

There is no broad consensus among criminologists and js researchers about whether optimism is warranted. Indeed, some researchers have been pessimistic about the potential payoffs of sophisticated, formal statistical approaches like those used in the present study.[9] Less sophisticated approaches have tended to perform about as well as, and sometimes even better than, the more sophisticated ones because of the joint limitations of weak theory and unavailable data. Without firm theory and proper data, it is virtually impossible to capitalize on the power of the more sophisticated statistical approaches. However, other researchers have been more optimistic about realizing the promise of sophisticated statistical analyses with respect to js applications, basing this optimism mainly on the fruits of their own

---

[9] Farrington, D. F., Tarling, R. 1985. Criminological prediction: the way forward. In <u>Prediction in Criminology</u>, ed. D. F. Farrington, R. Tarling, 258-68. Albany, NY: State University of New York Press; Gottfredson, S. D. 1987. Prediction: an overview of selected methodological issues. In <u>Prediction and Classification: Criminal Justice Decision Making, Special Issue of Crime and Justice, A Review of Research, Vol. 9</u>, 21-51. Chicago, IL: University of Chicago Press.

research enterprises.[10] But even these optimists have been circumspect in their expectations. There are limits to predictive knowledge inherent in both behavior itself and the research methods which can be mustered to explain behavior: all behavior entails some randomness and all research methods entail some error (e.g., sampling, measurement).[11]

Some useful prediction standards can be culled from recent reviews of the most rigorously conducted prediction research. One very broad touchstone suggests keeping both false positives and false negatives below 50 percent.[12] Another broad touchstone, focusing only on false positives, sets the somewhat less challenging sight of about 67 percent.[13] These rather high levels of predictive inaccuracy are not encouraging, and they take much of the wind out of the sails of all positions in the arguments about the ethical justifications of prediction applications. Arguments are weak, if not moot, when predictive inaccuracy is so great that no proponent would feel secure in advocating the formal adoption of these prediction instruments. If these levels of inaccuracy cannot be budged downward, all attempts at explicit and

---

[10] Schmidt, P., Witte, A. D. March 1987. Some thoughts on how and when to predict in criminal justice settings. Wellesley College, Wellesley, MA.: Department of Economics, pp. 26-27.

[11] Schmidt, P., Witte, A. D. March 1987. Some thoughts on how and when to predict in criminal justice settings. Wellesley College, Wellesley, MA.: Department of Economics, pp. 27-28.

[12] Farrington, D. P. 1987. Predicting individual crime rates. In Prediction and Classification: Criminal Justice Decision Making, Special Issue of Crime and Justice, A Review of Research, Vol. 9, 53-101. Chicago, IL: University of Chicago Press. Cited in Schmidt, P., Witte, A. D. March 1987. Some thoughts on how and when to predict in criminal justice settings. Wellesley College, Wellesley, MA.: Department of Economics, p. 26.

[13] Miller, M., Morris, N. 1988. Predictions of dangerousness: an argument for limited use. Violence and Victims 3,4:270.

formal prediction may be abandoned, and the js may stall and stagnate in its present posture of largely inexplicit and informal decision making.

While there is certainly some scientific cause to be optimistic, albeit cautiously so, about the future yield of prediction studies, we would continue to pursue work along these lines even if there was great cause to be pessimistic. We would be so inclined because the bulk of prediction research to date, whether it lends itself to optimism or to pessimism, has neither focused on serious violent behavior nor used the kinds of technical methods and analytical strategies employed in this study. That there is no glut of prediction research on criminal violence is as startling as it is bewildering. After all, violent crime probably represents for most people the defining and most disquieting aspect of the crime problem. Public consensus about this should certainly have spurred legislators to earmark public dollars for underwriting such research and motivated private researchers to respond to that challenge. But this has not been so. The considerable inertia and gaps in these respects makes most prediction research, because so broadly conceived and unfocused, only peripherally relevant to the present study.

SEQUENTIAL PREDICTION AND THE VIOLENT CRIMINAL CAREER

Arrests for violent crimes accumulate as sequential points in an individual's ongoing criminal career.[14] Front-line js practitioners need to make decisions about how to process a person arrested for a violent crime each

---

[14] Blumstein, A., Cohen, J., Roth, J. A., Visher, C. A., eds. 1986. Criminal Careers and "Career Criminals". Washington, DC: National Academy Press; Weiner, N. A. 1989. Violent criminal careers and "violent career criminals". In Violent Crime, Violent Criminals, ed. N. A. Weiner, M. E. Wolfgang, 35-138. Newbury Park, CA: Sage Publications, Inc.

time that person is arrested for such a crime. At each successive arrest, the decision maker must commonly take into account what has occurred earlier in the criminal career, including the criminal's own illegal behavior and js responses to that behavior, for instance, court convictions and institutional confinement. For this reason a <u>sequential-prediction</u> approach to decision making is useful: a prediction instrument needs to be developed at <u>each</u> point in the arrest sequence for use in deciding whether the arrested person will again be arrested for a violent crime.[15] Because some violent criminals accumulate many arrests, a chief goal of sequential prediction is parsimony, the creation of the fewest possible prediction instruments to cover as many points as possible in the arrest sequence. A first step toward parsimony is the separate examination of the prediction instruments devised at each point in the arrest sequence in order to see whether these instruments share common features amenable to consolidation. This strategy was adopted by the present study.

[15] This approach is distinct from <u>career-criminal prediction</u>. In career-criminal prediction, information compiled at the earliest points in the arrest sequence is used to predict which criminals will have <u>many</u> future arrests rather than, as is the focus of sequential prediction, at least one more arrest. Those characteristics of the violent criminal (e.g., age, gender, race) and the violent criminal career (e.g., number and types of prior arrests, convictions, and imprisonments) which signal at an early point in the career the accumulation of many future arrests for violent crimes may not be the same as those characteristics which signal at each arrest for a violent crime the accumulation of at least one more arrest for such a crime.

SOME MOTIVATING ASPECTS OF THE REPORT

Js decision making has three main aspects: goals, alternatives and information.[16] Goals are the aims, or objectives, which are sought by js officials through the decisions that they make. The more explicit these goals, the better equipped these officials will be to weigh and balance the relative importance of these goals and to assess competing strategies for their attainment. The present study takes the position that social utility, such as averting serious crimes either through sheer corporal incapacitation (physical restraint, not physical impairment) or through the delivery of effective social services, is probably the foremost js goal served by prediction-based classification. Alternatives are the choices, or decision pathways, from which the decision maker must select. Within the js, there are basically two kinds of decision making choices. The first kind, directly relevant to this study, pertains to deciding whether unlawful behavior is likely to recur in the form of rearrest: this kind of decision choice is risk prediction. The second kind of decision choice pertains to selecting a judicial disposition from the available options (e.g., confinement, intensive supervision, social-service provision) in response to the decision that the unlawful behavior is likely to recur: this is commonly called intervention (program) assignment. Information is the knowledge that some data are related and, therefore, relevant to specific decision making goals in the sense that this knowledge reduces decision making uncertainty. Information maximally

---

[16] Gottfredson, M. R., Gottfredson, D. M. 1988. Decision Making in Criminal Justice: Toward the Rational Exercise of Discretion, 2nd Edition. New York: Plenum Press.

represents a causal relationship between the data and the behavioral outcome; information minimally represents an associational relationship.

Motivated by the potential and desireable advancement in social utility made possible by prediction-based classification, the present study examined whether certain kinds of (mainly) official information about persons arrested for serious violent crimes might be relevant to decision making at selected early js decision points. In this respect, the study focused on only one kind of decision making choice, the extent to which selected information provided a basis for deciding that a particular behavioral outcome--rearrest for serious violence--was likely. The present study examined, then, risk prediction and, relatedly, prediction accuracy. The study did not look at whether the associated, subsequent decision to select a particular judicial disposition (i.e., an intervention or program), which might have been triggered by the predicted behavioral outcome, was a useful and effective decision (i.e., intervention or program evaluation). Whatever might be the diverse laudable or lamentable goals of a particular disposition with respect to a particular behavior, the capacity to reach that goal through that disposition will be impeded if the prediction instrument incorrectly identifies the actual behavioral outcome. Systematic intervention success, therefore, depends crucially on systematic prediction accuracy, and, for this reason, the assesment of predictive accuracy was a prominent concern of the present study.

These rather broad questions pertaining to js decision making reduced to some narrower ones. First, how accurate are time-related predictions of arrests for violent crimes? Second, is it practical to implement such predictions within the js? For example, Is predictively useful information readily available to key decision makers? Third, what can be done to improve prediction accuracy beyond whatever points are reached in this study?

These questions and analytical orientations are the report's motivations and cornerstones, and they are examined in a quite focused way. The present study focuses on the arrest sequence, linking the transition from arrest-to-arrest to decision making issues faced by the police, lower court judges, and supporting court personnel (e.g., probation officers). A systematic exploration of time-related prediction within the contexts of the juvenile and criminal justice systems naturally commences with the front-end arrest and related pretrial detention decision points. The present study focused on these front-end decision points. The present study further focused on those criminals who have been arrested for serious violent crimes. This behavioral focus is also a natural starting point for a systematic analysis of time-related prediction in view of the grave physical harm caused by, and the intense and widespread fear provoked by, violent criminals and their violent behavior. Future analyses of js processing points subsequent to arrest can potentially profit from whatever the insights, errors, and omissions of this front-end investigation; so too can analyses of nonviolent criminals.

This chapter began by spotlighting the persisting national patterns in sociodemographic concentrations of arrests for serious violent crimes in highly populated cities, and among young, minority males. Fortunately, enhancing the general relevance of this research, we have been able to study the members of two sizable male birth cohorts, comprising many minority members, as they matured through their juvenile and young adult years, and who resided for much of that time in one of our nation's most populated urban areas.

BREAKING SOME NEW GROUND: STUDY ADVANCEMENTS

The present study is the usual research hybrid, preserving continuites with past research by acknowledging and building upon the strengths of that research, and taking off in new research directions to see what horizons lay ahead. Several research decisions were prompted by some successful and challenging analytical applications in prior studies.[17]

This study took several steps to cement research continuities. First, the study examined almost all of the major time-related statistical models used in earlier js research, making sure to include the most promising of these models. This broad decision had two salutary conceptual and technical consequences: first, both "unitary-" (i.e., everyone is eventually rearrested) and "split-population" (i.e., only a segment of persons are eventually rearrested) statistical models were used, plugging us into some important conceptual currents which have become progressively more resonant in the criminological and criminal justice literatures; second, because the study employed most of the major statistical models currently under examination, it adopted the useful strategy of comparing statistical models exhibiting unidirectional ("monotonic") and multidirectional ("nonmonotonic") patterns in rearrest risks over time. Second, the study included in the statistical models many risk variables identified by some of the most powerful and recent studies as being related to recurrent criminal behavior. Third, the study explored the prospects of achieving highly accurate prediction-based classification with respect to both individuals and aggregates. Fourth,

---

[17] For a discussion of earlier research and the uses of time-related prediction, see Schmidt, P., Witte, A. D. 1988. Predicting Recidivism Using Survival Models. New York: Springer-Verlag.

repeating a quite common and useful practice, the study split the research subjects into two analytical segments, construction and validation groups, in order to assess the robustness and validity of results.

This study likewise stepped in some new research directions. First, the study focused only on those individuals who had been arrested for serious violent crimes, examining only their arrests for these serious crimes. It may startle the reader, but this is the first time (as far as we are aware) that a study has focused on just these criminals and just these crimes. Second, the study examined subjects in two sizable birth cohorts, enabling the assessment of the stability of results across time periods differing in selected social and historical respects. Third, capitalizing on the longitudinal nature of the birth cohort design, the study developed prediction-based classifications at each point in the arrest sequence in order to assess the stability of statistical findings and prediction accuracy across successive arrests. Fourth, the study compared both juvenile and adult arrest sequences, again capitalizing on the longitudinal structure of the birth cohort design. Fifth, the study mounted a substantial effort to locate as many subjects as possible in the two birth cohorts in order to ensure their internal representativeness, which, we expected, also enhanced their external representativeness, increasing the chances that the sequential-prediction models developed in this study might be capable of use with other groups.

Progress in developing useful prediction-based classifications lies in the measured combination of something old and something new when putting together one's research plan. We believe that such a balance has been struck here between earlier research advancements and new research technologies, making possible whatever might be this study's contributions to improving prediction accuracy.

An important caveat is worth emphasizing at this time. Although the study design punches beyond the envelope of past practices in several respects, and the statistical techniques that were used are among the most powerful of those now available, findings are nonetheless still preliminary. The study assessed only in a limited way the predictive potential yielded by the conjunction of an unusually large and rich data resource and statistical methods which are particularly well suited to exploiting that resource. Obviously, there are no ironclad guarantees that this opportune conjunction will lead to advancements in our capacity to predict who will be rearrested for serious violent crimes. The study was motivated by the simple desire to see whether this conjunction in data and methods might be cause for optimism that a practically defensible level of predictive accuracy might be achieved with respect to arrests for seriously violent crimes.

No single study can possibly answer either fully or definitively how well we can now predict or might in the future be able to predict recurrent arrests for serious violent behavior. That challenging objective amounts to a wide ranging and long-term research program requiring many different studies of many different places and times. However, a single study like this one can begin to assess the potential practical payoffs of traveling down certain methodological pathways.

ORGANIZATION OF THE FINAL REPORT

Chapter 2 describes the overall study design, including the samples, variables, statistical techniques, and validation procedures. Chapter 3 presents the final prediction models estimated for the juvenile and young adult periods. Chapter 4 summarizes the analyses and discusses next steps.

Chapter 2

METHODOLOGICAL ESSENTIALS: DATA SETS, RISK VARIABLES, RESEARCH DESIGN, AND STATISTICAL TECHNIQUES

THE STUDY DESIGN IN A NUTSHELL

The present study analyzed all arrests recorded by the Philadelphia police for involvements in violent index crimes (i.e., criminal homicides, forcible rapes, robberies, and aggravated assaults) by the black and white males in two birth cohorts, one born in 1945 and the other in 1958, who resided in Philadelphia from their tenth through their eighteenth birthdays and who were arrested at least once for a violent index crime sometime between their tenth and their twenty-seventh birthdays. Starting at a birth cohort subject's first arrest for a violent index crime, that first arrest and all subsequent arrests for violent index crimes were organized into an overall violent index-crime arrest history.[1] These arrest histories were then subdivided into those arrests which fell into (1) the juvenile years (ages ten through seventeen) or (2) the young adults years (ages eighteen through twenty-six). Individual and aggregate sequential-prediction analyses were separately conducted for each birth cohort, each age interval, and at each successive arrest in the violent arrest history.

Table 2.1 summarizes the above overview. The rows list the birth cohorts and the columns the numbers of birth cohort subjects who were at risk of arrest for violent crimes and the age intervals during which this risk was sustained. An "X" in a cell indicates that sequential-prediction analyses

---

[1] In this study, "violent index crime" is used interchangeably with "serious violent crime", "property index crime" with "serious property crime" and, more generally, "index crime" with "serious crime."

were conducted for subjects defined by the designated birth cohort and age interval.

Notice that the 1945 birth cohort appears twice: "total" and "follow-up". Arrest histories could be compiled for the full 1945 birth cohort for just the juvenile period and for a ten-percent adult follow-up sample for the young adult years. To bolster the reliability of results, the full 1945 birth cohort was used for analyses pertaining just to the juvenile period. The follow-up sample was used whenever age comparisons were made between arrests in the juvenile and the young adult periods or continuities in arrests were examined across the combined juvenile and young adult periods (thus, the inclusion of the age interval "10-26"). The text discussion and accompanying tables make clear whether the total 1945 birth cohort or the follow-up sample was used.

Table 2.2 presents the number of subjects in each birth cohort who were arrested for at least one violent index crime (<u>participants</u>), broken down by age interval and the number of arrests these subjects accumulated (<u>incidents</u>). The table also lists the number of arrest transitions (i.e., traversals from one arrest to the next) examined in the sequential-prediction analyses. As the table shows, the 1958 birth cohort subjects were divided into two subsets, a <u>construction</u> group (70 percent) and a <u>validation</u> group (the remaining 30 percent). These groups were basically used in a two-step procedure: first, sequential-prediction models, in the form of prediction equations, were developed using the larger construction group; second, to examine the robustness and generality of these models, they were applied to predictive decision making to the smaller validation group. (This split-sample design is a mainstay of prediction research.) The numbers of arrested subjects and their accumulated arrests are presented for the two incarnations of the 1958 birth cohort.

Clearly, the 1958 birth cohort produced the much larger number of arrested subjects and, partly because of this fact, also the much larger number of arrests. The total 1945 birth cohort, although less substantial in size, still generated sizable numbers of violent delinquents and violent arrests.

Although it was certainly sufficient for pursuing some complex and powerful statistical analyses, the total 1945 birth cohort was nevertheless more limited than the 1958 birth cohort. For example, the total 1945 birth cohort exhibited less than one-half the rate of participation of the 1958 birth cohort during the juvenile years (.036 versus .082), when both birth cohorts were at their total memberships.[2] And, as was pointed out, the total 1945 birth cohort was outproduced by the 1958 birth cohort with respect to the number of juvenile arrests for violent index crimes; scanning across the total 1945 birth cohort and just the construction group of the 1958 birth cohort, one calculates that the 1945 birth cohort generated three-fourths the rate of violent index crimes.[3] The follow-up sample of 1945 birth cohort understandably showed, because of the proportional sampling itself, even more modest numbers of both arrested subjects and, in turn, arrests. Because of the resulting disabled reliability, the follow-up sample was used only to

---

[2] The juvenile violent participation rate of the total 1945 birth cohort was 3.6 percent (360/9,945 = .036) in comparison to 8.2 percent for the total (construction plus validation groups) 1958 birth cohort (1,083/13,160 = .082).

[3] The total 1945 birth cohort produced 1.2 arrests per arrestee for violent index crimes (435/360); in contrast, the construction group of the 1958 birth cohort produced 1.6 arrests per arrestee for violent index crimes (1,655/759). In making these calculations, we used the number of arrests through the final arrest transition analyzed for each birth cohort.

assess the validity of the results obtained from analyses of the 1958 birth cohort construction group.

Table 2.3 shows for each birth cohort and age interval the number of birth cohort subjects who were arrested at each arrest transition. The table reinforces a quantitative point which by now is certain to be a commonplace: the 1958 birth cohort exhibited the much larger numbers of arrested subjects and arrests per arrested subject. For several reasons, the more abundant numbers of arrested persons and arrests recommended choosing the 1958 birth cohort as the departure point for developing the prediction instruments: (1) findings would be more <u>reliable</u>; (2) <u>more advanced points</u> in the arrest sequence could be examined; and (3) the <u>internal validity</u> of the prediction instruments could be scrutinized using a split-sample strategy.

The 1958 birth cohort was the natural choice as the departure point for these prediction analyses for another reason, in addition to its numerical superiority: because of its greater <u>recency</u>, results were more likely to have greater <u>currency</u>. The overall arrest history of a more contemporary birth cohort is more likely than that of a less contemporary birth cohort to reflect the overall arrest histories of birth cohorts which either are now passing through or will soon be passing through their juvenile and young adult years. The 1958 birth cohort matured through its juvenile and young adult years between 1968 and 1984, whereas the 1945 birth cohort did so 13 years earlier, between 1955 and 1971. More recent study data enhances the relevance of that data to current practical applications.

Using arrest- and judicial-history information obtained from police and court records, sequential-prediction analyses employing multivariate regression methods were conducted at each rung in the arrest sequence. Risk variables selected for these analyses overwhelmingly measured aspects of the official arrest histories of birth cohort subjects, up to and including their immediate arrests. Official information was emphasized for several reasons, which are discussed in the next section on data sources.

These are in a nutshell the main aspects of the study samples and research design. Together with the discussion in the first chapter, the reader is now equipped with the basic tools needed for tackling the next two analytical chapters. However, should the reader desire more detailed discussion of these issues, this detail is presented in the following sections of this chapter. These sections discuss selected characteristics of the birth cohort subjects and the segment arrested for serious violent crimes, describe data gathering procedures, define variables, and present the rationale for the study design and statistical techniques.

DATA SETS

The study used extensive information sifted from official police, court, and school records on 9,945 males born in 1945 and 13,160 males born in 1958 who resided in Philadelphia from their tenth through their eighteenth birthdays. To enhance the representativeness of the birth cohorts, a crucial asset to maximize when creating prediction instruments which are, one hopes, capable of application to other groups, the total populations of both birth cohorts were sought rather than sampled, and an intensive search was mounted to locate all persons who met the twin birthyear and residency requirements.[4]

---

[4] These efforts are detailed in Tracy, P. E., Wolfgang, M. E., Figlio, R. M. 1990. Delinquency Careers in Two Birth Cohorts. New York: Plenum Press; Wolfgang, M. E., Figlio, R. M., Sellin, T. 1972. Delinquency in a Birth Cohort. Chicago, IL: University of Chicago Press.

The decision to include all birth cohort subjects, accompanied by our vigorous attempt to locate the records of as many of these subjects as possible, undoubtedly bolstered the internal representativeness of the final complements of birth cohort subjects, increasing the reliability of the analyses. But these research decisions and procedures were also important because they helped to bolster the external representativeness of the two birth cohorts and, by so doing, to increase the generality of the findings.

The prediction instruments developed here using the two birth cohorts can be effectively applied to other groups to the extent that these two cohorts share common violence-related characteristics and causal processes with these other groups. This point is, of course, only to reiterate that the birth cohorts should display as much external representativeness as possible. But, however much one might try to bolster external representativeness, one can still never be entirely certain that the prediction instruments can be applied to other groups. This uncertainty exists for two reasons. First, while one is conducting a prediction study, one can never fully anticipate which groups will be selected for later applications. Consequently, the degree of external representativeness is a matter mainly to be determined on a case-by-case basis, each time the prediction instrument is applied to a new group. However, the immediate generality of one's study group can be buttressed by wise sampling decisions, based on one's vision of how the study findings are most likely to be applied in the near future. Second, a study's external representativeness is not static but dynamic, changing over time in response to changes in the violence-related characteristics and causal dynamics of groups. Obviously, this dynamic aspect of representativeness is not peculiar to prediction research but is present in all causal and correlational research. One way to handle this particular problem, the natural decay of prediction findings, is periodically to update one's prediction instruments, doing so whenever predictive accuracy falls below some prespecified, unacceptable threshold.

The residency requirement--uninterrupted residence in Philadelphia from age 10 to 18--helped to ensure that members of both birth cohorts experienced, in common, major socioeconomic conditions and js operations and policies while passing through their juvenile years. In the language of research design, the residency requirement aided in "controlling for," or "equalizing" across subjects within each birth cohort, broad socioeconomic and js influences on serious violent criminal behavior, so that the effects of other personal and criminal history variables might be more sharply and accurately detected. The residency requirement also helped to ensure that the members of both birth cohorts remained in the city sufficiently long to generate citywide participation and arrest rates which were large enough to sustain reliable analyses. This rather stringent requirement was also prompted because of the administrative obstacles commonly faced when trying to gather police records from several jurisdictions. The cost of gathering such records were also prohibitive. Tracy, Wolfgang, and Figlio (1990) have described how the two birth cohorts were identified, their sociodemographic compositions, the rationales for selecting the cohort birthyears, procedures followed to collect and code official records, and the diverse information gathered about personal, social, and crime characteristics.[5] Some basic information about these issues is presented below.

---

[5] Tracy, P. E., Wolfgang, M. E., Figlio, R. M. 1990. Delinqency Careers in Two Birth Cohorts. New York: Plenum Press.

The research program studying the 1945 birth cohort subjects generated a "total" cohort, which was tracked across the juvenile years, and a more modest sized "follow-up" sample drawn from the cohort, which was tracked across the young adult years. It was methodologically unnecessary in view of commonly accepted sampling theory and, at the time these data were being collected, prohibitively expensive to study the arrest chains of all 9,945 cohort subjects as they aged beyond their juvenile statuses. Consequently, a ten-percent stratified random sample was drawn from the total birth cohort. This is what we have called the "follow-up" sample.

The research program studying the 1958 birth cohort subjects generated a single "total" birth cohort. Consistent with the earlier research program which looked at the 1945 birth cohort subjects, it was still methodologically unnecessary to study the arrest histories of all 13,160 birth cohort subjects as they aged into adulthood. However, because of the great strides in reducing computing costs and time, it was no longer prohibitively expensive to do so. For this reason, a single, total 1958 birth cohort was used for all analyses. Greater analytical reliability and moderate analytical costs presented an unusual research opportunity which was promoptly seized.

Arrest and judicial histories were compiled from police and court records for all birth cohort subjects who had been arrested for at least one violent index crime between the ages of ten to seventeen (the juvenile period) or eighteen to twenty-six (the young adult period), reflecting the juncture at which the Pennsylvania js splits into its juvenile and adult tiers.[6]

---

[6] We used age twenty-six as the upper age limit because arrest records were unavailable beyond this point for the 1958 birth cohort subjects. To ensure that we used a common age range in the two cohorts, twenty-six was also used as the upper age limit for the 1945 follow-up birth cohort. Fortunately, given our research purposes, the great bulk of arrests for serious crimes, in

Individual arrest sequences were then generated for these subjects based on all of their violent index crimes. These arrest sequences formed the bases for both the individual- and aggregate-level sequential-prediction analyses.

Official Records

Official records were exclusively used for two reasons. First, the prediction instruments which were developed might possibly be used by js officials at the time an arrest is made. The chances for such use would be greater if the prediction instruments incorporated <u>risk (predictor) variables</u> which, in addition to their requisite predictive capacities, were easily accessible and quickly available to key decision makers. Official information was attractive for this reason. It is true that the incorporation of unofficial information, obtained from survey instruments (perhaps short screening interviews or questionnaires) administered to the arrested person, might have resulted in more powerful prediction instruments. It is also true, however, that these instruments would almost certainly have been more cumbersome routinely to use, and, because of this, they would have been that much less likely to be adopted by administrators and key decision makers. Second, the <u>outcome (predicted) variable</u> was the "time until rearrest." Official police and court records are the most reliable sources of information about this timing.

One might question using an official measure, rearrest for serious violent crimes, as the outcome variable because it is not a measure only of

---

these birth cohorts and in others, is concentrated at the older juvenile and young adult ages. The substantive effect of this upper age barrier seems, therefore, to be small.

the individual's behavior but rather of the convergence of decisions made by several parties--by the arrested criminal (e.g., who to victimize, where to victimize that person), by js officials (e.g., the level of resources to devote to detecting crimes and apprehending criminals), and by the victims (e.g., whether to report the crime to officials). Under some circumstances, one might certainly choose to predict the recurrence of serious violent behavior rather than arrests for that behavior. This would be the preferred choice if one were interested in tracing the causal roots of the violent behavior itself. But, if one is interested in the organizational implications of this behavior, manifested as an arrested person who must be processed by js officials, then the rearrest measure is the appropriate outcome variable. Predicting an arrest for serious violent behavior with some precision can aid in targeting which and/or how many persons will enter the queue of clients to be processed by the js. The study is chiefly concerned with this organizational aspect of seriously violent behavior.

a. The Juvenile Years

The Juvenile Aid Division (JAD) of the Philadelphia Police Department handles all police contacts by juveniles resulting in arrest or diversion to a social service agency ("remediation"). The JAD maintains an up-to-date, hard-copy summary of all police contacts for each juvenile apprehended for a delinquent act--the widely referred to "rap sheet." For the 1945 birth cohort, these records covered the years 1955 to 1963, and for the 1958 birth cohort, these records covered the years 1968 to 1976.

The rap sheet is quite useful as a research aid because it briefly catalogues the juvenile's delinquent career and provides some demographic description of the youngster. However, the rap sheet does not present a complete and detailed account of each police-contact incident; it simply indicates the date on which the incident occurred, the police district in which the incident took place, and the complaint number assigned by the JAD to the incident, which can be used to track the incident through the dispositional vicissitudes of the juvenile justice system.

The complaint number appearing on the rap sheet was used to locate the hard-copy police Investigation Report, which provides a much more detailed description of the police contact: where the incident took place; a demographic profile of the complainant; the number, genders, ages, and races/ethnicities of persons other than the cohort subject who participated in the incident; the type and extent of physical injuries and property losses sustained by victims; whether a weapon was used and, if so, the type of weapon; whether alcohol or other drugs were detected present in the incident; and the initial court disposition which was rendered. Some additional information was obtained from the police Arrest Report. The police Arrest Report, quite obviously, lists important arrest-related information: the time, date, and place of arrest; the number, genders, ages, and races/ethnicities of those arrested; and the official crime-code classification assigned by the police to the incident for which the arrest was made.

The police Investigation Report was the chief source of information used to create study variables. Some information was used just as it appeared on the report. Other information was statistically reworked into scales designed to calibrate the seriousness of the incident.

Some key biographical information was obtained from school records. For example, the birth cohort subject's race was obtained from these records because these records more accurately identify this information than do police records. The birth cohort subject's socioeconomic status (SES) was also derived from school records. The birth cohort subject's home address was used to identify the census tract in which the youth resided, and selected SES data corresponding that census tract were then used to measure SES level.[7] SES was, therefore, an aggregate spatial measure assigned to the birth cohort subject based on the subject's residential address. The SES measure calculated during the juvenile period was also applied to the young adult years.[8]

Philadelphia Family Court records were used to obtain information on judicial dispositions imposed for the birth cohort subjects' violent criminal involvements. One of the most important pieces of information obtained from these records was whether there was an affirmative adjudication of guilt and, if there was one, the kind disposition which was imposed.

---

[7] In the 1945 birth-cohort study, SES was measured as a five- category ordinal variable based on the median income level in the census tract in which the birth-cohort subject resided. These categories ranged, in ascending order, from poverty, deprivation, semideprivation, modest-but-adequate, to comfort. (For more details see Wolfgang, M. E., Figlio, R. M., Sellin, T. 1972. Delinquency in a Birth Cohort. Chicago: University of Chicago Press.) In the 1958 birth-cohort study, SES was measured as a continuous variable based on a principal components analysis of ten census-tract level measures of SES (e.g., income, education). (For more details about this procedure, see Weiner, N. A. 1986. Violent recidivism among the 1958 Philadelphia birth cohort boys. Final report submitted to the National Institute of Justice. Sellin Center for Studies in Criminology and Criminal Law: Philadelphia, Pennsylvania.)

[8] We were forced to do so because SES measures applying to the young adult years were not available at the time these analyses were done. This procedure makes one, or both, of the following assumptions: (1) the birth cohort subject's SES level was stable across the juvenile and young adult periods and/or (2) the birth cohort subject's SES level during the juvenile years exerted an influence on violent criminal behavior extending into young adulthood.

b. The Young Adult Years

The procedures used to collect adult arrest records for the 1945 birth cohort subjects through their 27th birthdays (from 1963 to 1972) basically mirrored the procedures used to collect their juvenile arrest records. Hard-copy police rap sheets were gathered and searched to identify and locate police Investigation and Arrest Reports, which were nearly identical in content to those described earlier in the discussion of the JAD records. Information obtained from these two police report forms was then coded similarly to the juvenile information.

Since 1971, summary information about adult arrests in Philadelphia have been entered onto computer files maintained by the city's Court of Common Pleas. These files, which are the computer-copy equivalents of the hard-copy rap sheets, were used to locate the more detailed police Investigation and Arrest Reports of the 1958 birth cohort subjects, who advanced through young adulthood between the years 1976 and 1985, well after the computerized record system had been installed. Information from these reports was compiled and then coded in the same manner as the information from their juvenile counterparts. The coding procedure replicated the one used for the 1945 birth cohort subjects.

For both the 1945 and the 1958 birth cohorts, complete adult arrest histories were merged with complete delinquent arrest histories, yielding continuous arrest histories from age ten through age twenty-six for each birth cohort subject. Based on these arrest histories, subjects arrested for violent index crimes were identified. And, as previously noted, individual arrest histories were then generated for these subjects, commencing with their first arrests for violent index crimes and terminating with their last arrests for these crimes.

SELECTION OF BIRTH COHORT SUBJECTS: GENDER AND RACIAL RESTRICTIONS

Only males were selected for study. Both artifact and substance prompted this limited focus. From its very inception, the 1945 birth cohort was so restricted in gender.[9] The present study must obviously conform to the methodological rule imposed by that original research decision. The 1958 birth cohort study was not so restricted in gender, representing new thinking on the part of the original research team. However, as it turned out, too few females in the 1958 birth cohort accumulated enough arrests for serious violent crimes to underwrite reliable analyses: during their juvenile years, only 140 females, from among 14,000 subjects, were arrested one or more times for a violent index crime; only 13 were arrested two or more times for such crimes; only 4 were arrested three times; and none were arrested more than three times.[10] This arrest pattern discourages even moderately complex statistical analyses of violent criminal careers. Reluctantly, the female birth cohort subjects were excluded from the study.

A similar impoverishment also forced us to restrict the study to blacks and whites. Among the 13,160 male birth cohort subjects, there were just 122

---

[9] The rationale for this selection strategy can be found in Wolfgang, M. E., Figlio, R. M., Sellin, T. 1972. Delinquency in a Birth Cohort. Chicago, IL: University of Chicago Press.

[10] Weiner, N. A. 1986. Violent Recidivism among the 1958 Philadelphia Birth Cohort Boys. Report to the National Institute of Justice, Center for the Study of Crime Correlates and Criminal Behavior. Philadelphia, PA: University of Pennsylvania, Sellin Center for Studies in Crimiology and Criminal Law. Appendix 2.

Hispanics, 6 Native Americans, and 4 Asian Americans.[11] Only twenty-eight of these subjects were arrested one or more times for a violent index crime, not nearly enough to permit solid analyses.

VARIABLES

Serious Violent Crimes and Serious Violent Criminals

If a birth cohort subject was arrested for a serious violent crime, that subject was included in the study. This selection rule might appear quite easy to apply in practice: for each arrested subject, simply scan each one of that subject's arrests to determine whether at least one of them was for a serious violent crime--a homicide, rape, robbery, or an aggravated assault--and, if one was, place that arrest and all subsequent arrests for serious violent crimes into a continuous violent-crime arrest history.

This selection rule may be quite simple in description, but it is not simple in fact. Classifying an arrest with respect to its crime type is often a difficult and confusing operation because a single incident can involve several criminal behaviors. For example, an attack ending in a fatality is both an aggravated assault and a homicide. If the attack involves forcible sexual intercourse, then it is a rape as well. The more criminal behaviors entailed by the incident, the more behaviorally dense the incident and, in turn, the more complex the classification task. This task can, however, be made more tractable by adopting some conventional crime-classification rules

[11] Weiner, N. A. 1986. Violent Recidivism among the 1958 Philadelphia Birth Cohort Boys. Report to the National Institute of Justice, Center for the Study of Crime Correlates and Criminal Behavior. Philadelphia, PA: University of Pennsylvania, Sellin Center for Studies in Crimiology and Criminal Law. Appendix 2.

which, although not entirely satisfactory, have substantial practical and, one hopes, theoretical utility.

In Philadelphia, as in most jurisdictions, a unique numerical code is assigned to each unlawful act defined in the State Criminal Statutes; the lower the crime code number, the more behaviorally grave the crime and, mainly because of this, also the more legally serious the crime. In the present study, arrests for violent index crimes were defined in terms of a <u>single</u> violent crime type, based on the lowest and, thus, most <u>serious</u> crime code assigned to that arrest. A lower crime code had priority over a higher crime code. This hierarchical rule enabled the assignment of a unique violent index-crime code to each arrest. The definitional hierarchy, from lowest violent index-crime code to highest violent index-crime code was as follows: homicide, rape, robbery, and aggravated assault.

Clearly one of the drawbacks of using only statutory codes to characterize arrests, especially a single crime-code designation when several might actually apply, is that the <u>behavioral</u> complexity of the crime can be clouded. Robbery perhaps most clearly illustrates this point. As conventionally defined by the Federal Bureau of Investigation's Uniform Crime Reporting System, robbery involves the threatened, attempted, or completed application of physical force to obtain something of value from another person against that person's will. But as the definition fully acknowledges, physical force need not actually be used. Thus, moreso than the other serious crimes, robbery perhaps poses the greatest threat to valid crime classification because robbery crime codes commonly fail to indicate whether force was used and, furthermore, whether physical injury was inflicted on a victim. Adding to this classification fog, the crime code may not clearly indicate whether something of value was actually taken during the robbery.

To date, no crime-classification system has been devised which entirely solves the vexing problem just outlined--boiling down the discrete multiple criminal behaviors, which one might think of as dimensions or components, comprising a crime incident into a single summary measure. However, one can augment, although probably not completely dispense with, the hierarchical classification rule based on crime codes and, by so doing, lessen its potential masking effect. By so doing, one can forge a workable overall classification protocol. This can be done simply by recognizing that the various crime codes reflect behavioral components of the criminal incident and that these behavioral components, in turn, reflect (among other things) the seriousness of that incident. In keeping with this line of reasoning, a seriousness measure was adopted by this study which summarized in a single score, falling on a ratio scale, the crime incident's spectrum of harmful behavioral components, as described in the narrative on the police Investigation and Arrest Reports.[12] If there is not an exact correspondence

---

[12]. Weiner, N. A. 1986. Violent Recidivism among the 1958 Philadelphia Birth Cohort Boys. Report to the National Institute of Justice, Center for the Study of Crime Correlates and Criminal Behavior. Philadelphia, PA: University of Pennsylvania, Sellin Center for Studies in Crimiology and Criminal Law. Appendix 7 describes the construction and content of the seriousness scoring scale. The main aspects of scoring are: the degree of physical injury inflicted or medical attention required, the amount of property theft and/or damage, the presence of forcible sex, the type of personal threat or intimidation, premises forcibly entered, and motor vehicle theft. Further details of the scale construction and rationale can be found in Sellin, T., Wolfgang, M. E. The Measurement of Delinquency. 1978. Reprint. Montclair, N.J.: Patterson Smith. A revised version of the seriousness scoring scale, based on a national sample, can be found in Wolfgang, M. E., Figlio, R. M., Tracy, P. E., Singer, S. I. 1985. The National Survey of Crime Severity. Washington, DC: U.S. Government Printing Office.

between the crime-code classification and seriousness-score procedure, the seriousness score can potentially register behavioral components of the criminal incident which the grosser crime code classification misses, <u>when used in conjunction with that grosser measurement</u>. The seriousness score is also useful in summarizing an individual's delinquent or criminal history, in the form of a mean, or average, seriousness score.

<u>The Playlist: Types and Selection of Risk Variables</u>

This study mainly searched for risk variables which were predictively related to a high probability of rapid rearrest for serious violence. The search was, however, constrained by some considerations beyond the usual commonplaces of time, money, and the absence of clear theoretical and empirical signposts: <u>ethicolegal</u> and <u>administrative</u>. The ethicolegal constraints were particularly knotty and centered on some sensitive and volatile, legal and political issues relating to the propriety of explicitly using certain classification criteria (i.e., measured as variables) in formal predictive decision making in the juvenile and criminal jss, resulting in the differential processing of arrested criminals (e.g., preventive detention versus intensive supervision).

Surprisingly, there are virtually no controlling statutory nor constitutional doctrines prohibiting the use of prediction and classification systems in cj decision making.[13] This legal vacuum is all the more surprising

---

[13] The following discussion follows closely Tonry, M. 1987. Prediction and classification: legal and ethical issues. In <u>Prediction and Classification: Criminal Justice Decision Making, Special Issue of Crime and Justice, A Review of Research, Vol. 9</u>, ed. D. M. Gottfredson, M. Tonry, 367-413. Chicago, IL: University of Chicago Press.

when one considers the potentially momentous impacts of this decision making, especially insofar as it relates to predictions of future criminal dangerousness and, in turn, the decisions about whether to impose, for example, pretrial (i.e., preventive) detention and to initiate priority prosecution. One might mount challenges against the use of certain classification criteria by invoking the constitutional principles of due process, equal protection, and cruel and unusual punishment, but such challenges have rarely been pursued. These objectionable classification criteria, widely spurned in other contexts as constitutionally noxious, such as in employment litigation and capital case processing, include race and ethnicity, religion, political affiliation, and possibly gender. However, with the exception of these classification criteria, there are precious few others which might be sufficiently constitutionally offensive to compel their judicial rejection as acceptable decision making criteria. Because the constitution is resoundingly mute about the nature of these criteria, ethicolegal considerations reduce mainly to ethical and public policy considerations.

The ethical questions centered on the propriety of using certain variables in js decision making when these variables do not focus on a persons's <u>behavior</u> but rather characterize either a person's <u>status</u> or personal attributes which lay <u>beyond a person's control</u>. Probably most noteworthy among the status variables is the socioeconomic triad: occupation, income, and education. And, probably most noteworthy (indeed, notorious) among the variables beyond a person's control is the demographic triad: race and ethnicity, gender, and age. These variables--but most notably race and ethnicity--have shaky legal standings as elements of formal js decision making because they represent intrinsic qualities of a person rather than unlawful extrinsic behavior (and consequences of that behavior) which a person has willfully performed.

It is entirely possible that both status characteristics and characteristics beyond a person's control may have substantial power as predictors. However, these characteristics, when transformed into predictor variables, may not pass legal muster as proper components of decision making instruments; their use violates the juridical principle that proper js decision making should be based on the assessment of a person's blameworthiness, or criminal intent (the mens rea principle), as inferred from freely chosen unlawful behavior (the actus rea principle) rather than from a person's preexisting social placement or, more importantly, personal qualities which cannot be altered through intentional behavior. Both kinds of characteristics have no direct relationshp to blameworthiness and, because of this, affront deeply rooted legal and social principles concerning the conditions under which criminal liability can justly be assigned. Variables which are legally prohibited or strictly limited in their judicial application are known as "suspect classes."[14]

Administrative concerns mostly centered on practical questions, such as how easily and routinely to provide information to decision makers that is, in turn, easily and routinely useful to them. Even the most powerfully predictive information will be reduced to a decision making dud if that information cannot be quickly and reliably compiled for prompt dissemination

---

[14] Cohen, J. 1983. Incapacitation as a strategy for crime control: possibilities and pitfalls." In Crime and Justice: An Annual Review of Research, vol. 5, ed. M. Tonry, N. Morris, 1-84. Chicago: University of Chicago Press.

to front line js decision makers. The best information is usually in the form of a short, crisp list of risk variables, accompanied by clear instructions for their use.

Variable selection was chiefly guided in tandem by ethicolegal and administrative considerations, and the study design was created to accommodate these concerns. As it turned out, the great majority of variables which were selected for examination are actually quite accessible to js officials and pose little administrative impediment to timely decision making. Fortunately as well, many of these variables also passed research muster: related studies on criminal careers indicated some predictive capacity on their part.[15] The ethicolegal issues, on the other hand, posed somewhat greater demands on the study because they impelled that we justify on ethicolegal grounds those variables which might be selected for examination. We now turn to that justification.

As we have reiterated, predictive decision making is at the very heart of the js. However, even the most accurate predictive decision making has a naked, cold edge unless it is draped in purpose; js decision making must be principled; that is, it must be rationalized. Two polar positions have emerged as the high grounds in debates about the proper basis upon which to make decisions about how to process arrested criminals: just (commensurate) deserts (i.e., retribution) and utilitarian incapacitation (e.g., selective and collective). Cohen and Tonry have carefully elaborated these positions:

---

[15] These variables include, among others, the prior individual crime rate, the type of first crime, and the age at first criminal involvement. For a comprehensive review of the research literature on predicting individual crime rates, see Farrington, D. P. 1987. Predicting individual crime rates. In Prediction and Classification: Criminal Justice Decision Making, Special Issue of Crime and Justice, A Review of Research, Vol. 9, ed. D. M. Gottfredson, M. Tonry, 53-101. Chicago, IL: University of Chicago Press.

their substance, friction points, and compatibilities.[16] The following brief comments have drawn heavily from their discussions.

In its widest institutional interpretation, regardless of the decision point in question, just deserts asserts that the choice of a js disposition, which ultimately is a quest to ascertain the proper severity or intensity of that disposition, depends on the amount of harm that a criminal inflicts on the victim and on the degree of culpability of that criminal. Strictly interpreted, only these two features of the immediate criminal incident ought to be explicitly used in formal decision making. Essentially, dispositions are selected because they are inherently, morally and legally "deserved," not because they serve some useful purpose. The js exerts its authority and, ultimately, power to right the criminal's wrong in a number of ways, by publically censuring the criminal and, thereby, solidifying social cohesion and by resetting the moral equilibrium, upset by the criminal's behavior, through condemnation and punishment. By doing these things, the state restores the moral balance and compass in society, disturbed by a criminal incident. The attainment of this end is a chief basis for asserting that the imposed disposition was justly deserved.

A modified version of this position relaxes the requirement that only aspects of the immediate criminal incident can be considered in decision making. The number and gravity of prior crimes resulting in juvenile

---

[16] Cohen, J. 1983. Incapacitation as a strategy for crime control: possibilities and pitfalls." In Crime and Justice: An Annual Review of Research, vol. 5, ed. M. Tonry, N. Morris, 1-84. Chicago: University of Chicago Press; Tonry, M. 1987. Prediction and classification: legal and ethical issues. In Prediction and Classification: Criminal Justice Decision Making, Special Issue of Crime and Justice, A Review of Research, Vol. 9, ed. D. M. Gottfredson, M. Tonry, 367-413. Chicago, IL: University of Chicago Press.

adjudications and adult convictions might also find a home in the decision making machanism, but these aspects of prior crimes would be given limited weight relative to aspects of the immediate crime.[17] Whether the strict or modified stance is assumed, just deserts is driven forward by the past: decision making choices are contingent upon the character of the criminal infraction which has just been committed and, perhaps, upon the judicial decisions which have been made about previous crimes. The past overwhelmingly informs present decision making.

Utilitarian incapacitation, on the other hand, is driven forward by the future. This position acknowledges the cold fact that persons who are placed in secure confinement cannot participate, while so confined, in serious crimes in civil society. Expectations of enhanced public protection justify the <u>differential imposition</u> of js dispositions and, given their imposition, the <u>differential harshness</u> of these dispositions. The approach is guided by the principle that preventive public protection, and related considerations of the economic efficiency and social effectiveness of such protection, is an ethically sound basis for choosing dispositions to impose on different persons; disparate treatment of persons who have committed the same type of crime is justified if these persons are judged to pose different risks of future criminal involvement. The decision making weights of the immediate

---

[17] Prior criminal record burdens the criminal during present js decision making because it reflects bad character and demonstrates wickedness, contempt for the law, the failure of past leniency, and misplaced earlier benefits of the doubt. For further discussion, see Tonry, M. 1987. Prediction and classification: legal and ethical issues. In <u>Prediction and Classification: Criminal Justice Decision Making, Special Issue of Crime and Justice, A Review of Research, Vol. 9</u>, ed. D. M. Gottfredson, M. Tonry, 367-413. Chicago, IL: University of Chicago Press.

crime and, moreso, prior crimes may pale in comparison to the weights given to future frequent and serious criminal behavior which might be foregone by an appropriate and timely, immediate js disposition. The utility of the decision is broadly gauged, then, by the number of serious crimes which are averted. In short, the capacity to attain future ends by present decisions overwhelmingly guides present decision making.

While these two legitimations of js decision making may seem to represent antagonistic and irreconcilable polarities, they are not necessarily so. There need be no final showdown. A midway accommodation acknowledges that the harm inflicted by a criminal at the time of the immediate crime, coupled with the criminal's blameworthiness at that time, are the chief criteria to use in making case-processing decisions. Within the range of dispositions stipulated by criminal statutes--which theoretically are based on just deserts precepts--incapacitation principles can be applied. Under this construal, just deserts standards set the <u>limits</u> within which incapacitation decisions must operate: officials can impose unlike dispositions upon criminals who have inflicted identical amounts of harm and who have exhibited the same degree of blameworthiness if these criminals pose different risks of future criminal involvements. However, this disparity is permissible only if it comports with the strict proviso that the imposed dispositions must all fall within the range of dispositions fixed by just deserts principles; that is, the imposed disposition must not be undeserved, and this desert is guaranteed because it is fixed by law.

The midway accomodation forms the framework for the variable selection and study design of the present research. Risk variables endorsed by the just deserts formulation became the foundation of these analyses. These risk variables mainly included aspects of the immediate serious crime. However, in acknowledgment of the midway position described above, some aspects of a subject's prior arrest history were also employed, such as the number and gravity of those prior crimes for which the subject was adjudicated or convicted. If decision making variables assembled under the just deserts banner were also to prove useful in risk assessment, then these variables would have the unexpected but salutary secondary payoff of promoting incapacitation objectives. Suspect classes would be automatically unacceptable for direct, explicit application even if they served to further incapacitation goals. However, suspect classes (e.g., race) did have a legitimate and important role in identifying risk variables which are ethically proper to include in a prediction <u>instrument</u> but which are ethically improper to include in the prediction <u>application</u>. (More will be said about this later.)

The variable selection and study design were informed by these ethicolegal and administrative concerns, and by the constraints of principled js decision making. A battery of risk variables was selected based upon the joint considerations of their ethicolegal propriety and administrative utility. As it turns out, these risk variables had the added bonus of being among those variables commonly spotlighted by both public speculation and scientific research.[18]

---

[18] For a comprehensive review of research on predicting individual crime rates and of risk variables generally found useful in this respect, see Farrington, D. P. 1987. Predicting individual crime rates. In <u>Prediction and Classification: Criminal Justice Decision Making, Special Issue of Crime and Justice, A Review of Research, Vol. 9</u>, ed. D. M. Gottfredson, M. Tonry, 53-101. Chicago, IL: University of Chicago Press.

We first restricted the analysis to those risk variables characterizing the immediate arrest. These variables have been designated the ethically and legally <u>permissible</u> subset and represent the overall harmfulness of the incident which resulted in an arrest. We then widened the analysis to include risk variables characterizing the subject's prior criminal record and selected sociodemographic attributes. These risk variables have been designated the ethically and legally <u>less permissible and impermissible</u> subset.

The research strategy based on this variable characterization took the following general form. First, we determined whether the ethically and legally permissible risk variables were predictively related to the timing of rearrests for serious violent crimes; second, we determined whether both the ethically and legally less permissible risk variables and the outright legally impermissible risk variables were related to the timing of rearrest. While both the ethically and legally less permissible variables and the grossly legally impermissible variables might contravene standards for principled judicial decision making, they might nevertheless possess predictive and conceptual significance and, thus, have deserved examination on purely intellectual grounds. But these less permissible and clearly impermissible variables were also worth examining for ethicolegal reasons: as other researchers have properly argued, a known suspect class <u>must</u> be explicitly included in the <u>initially estimated</u> prediction instrument in order to purge that variable's influence from the <u>finally applied</u> prediction instrument.[19]

The purging procedure involved two phases as we moved from prediction-model estimation to prediction-model application: first, the impact of the

---

[19] Schmidt, P., Witte, A. D. March 1987. Some thoughts on how and when to predict in criminal justice settings, working Paper. Wellesley College, Wellesley, MA: Department of Economics, p. 31.

objectionable variable was explicitly accounted for in the initial estimation stage by the inclusion of that variable at this stage; second, once the effect of the objectionable variable had been estimated and accounted for, that variable was then omitted from the predictive decision making instrument, thereby neutralizing its effect in the applied setting. Although omitting these variables from the applied setting may diminish predictive accuracy, doing so nevertheless helps to ensure that these variables do not indirectly, inappropriately influence the prediction instrument through a backdoor association with some other included ethically and legally permissible variable. It remains to be seen, however, whether predictive accuracy hinges to any great extent on these suspect variables.

Based on the above considerations and prior research results, two sets of risk variables were created. The first set comprised the ethically and legally permissible risk variables, based on a strict just deserts interpretation of permissibility. The second set comprised both the ethically and legally less permissible risk variables and the outright legally impermissible risk variables. The ethically and legally less permissible variables in the second risk-variable set ranged from aspects of the prior criminal history (which mostly verged on ethical and legal propriety) to sociodemographic attributes (which mostly verged on or fell into ethical and legal impropriety).

Each prediction model in the arrest sequence was then estimated as follows: First, all of the ethically and legally permissible risk variables (in Set I) were forced into the sequential-prediction model along with race (in Set II), the most notorious suspect class, in order to purge its effect. This analysis enabled us to determine the overall predictive value of the ethically and legally permissible variables, net the effect of race. Second, all of the ethically and legally questionable risk variables (in Set II) were forced into the sequential-prediction model along with all of the legally permissible risk variables (in Set I) and the race variable (again to purge its effect). Using this strategy, we were able to determine whether the ethically and legally permissible risk variables held any predictive value, net the impact of race, and, further, whether the ethically and legally less permissible risk variables contributed anything more of predictive value. Figure 2.1 presents the battery of risk variables, the risk-variable set into which the variables fell, and how the variables were measured.

RESEARCH DESIGN AND STATISTICAL TECHNIQUES

Viewing Rearrest For Serious Violence as a Failure Time Process

The family of statistical techniques used in this study is generally called failure time analysis or, alternatively, survival analysis. What one calls these techniques depends upon the field of application and the way one conceives of the phenomena studied in these fields. Engineering uses these techniques to study system failure (e.g., electrical, mechanical, or structural breakdown), whereas biomedicine and epidemiology uses them to study organismic survival (e.g., remission following physical trauma or exposure to disease or toxic agents).[20] We prefer the term "failure time analysis"

---

[20] The most general term for this family of techniques is "event history analysis." For detailed discussions of the techniques, one can consult: Allison, P. A. 1982. Discrete-time methods for the analysis of event histories. In Sociological Methodology, 1982, ed. S. Leinhardt, 61-98. San Francisco: Jossey-Bass; Allison, P. A. 1985. Event History Analysis. Beverly Hills: Sage Publications; Cox, D. R., Lewis, P. A. W. 1966. The Statistical Analysis of Series of Events. London: Methuen; Cox, D. R., Oakes,

because it aptly reflects the idea that rearrest for a serious violent crime represents the failure either (1) to refrain from involvement in one of these crimes or, less optimistically, (2) to avoid detection and arrest for involvement in one of these crimes. Strictly speaking, the present study addressed only the second of these two meanings: "failure" only signifies that the birth cohort subject was unable to remain free from arrest during the observation period.

Failure time techniques are statistical methods for analyzing the structure of random variables which can take on only positive values, such as the time interval from arrest to rearrest for serious violent crimes. These techniques represent an important advancement beyond the traditional prediction techniques which defined rearrest simply as the occurrence of at least one more arrest within a fixed follow-up period (commonly eighteen months to thirty-six months). Several researchers have noted the deficiencies of these traditional prediction techniques, collectively called the "binomial" approach to studying rearrest risks.[21]

D. 1984. The Analysis of Survival Data. New York: Methuen; Holden, R. T. 1983. Failure time models for criminal recidivism. Unpublished paper. New Haven, CT: Yale University, Department of Sociology; Kalbfleisch, J. D., Prentice, R. L. 1980. The Statistical Analysis of Failure Time Data. New York: Wiley; Lawless, J. F. 1982. Statistical Models and Methods for Lifetime Data. New York: Wiley; Lee, E. 1980. Statistical Methods for Survival Data Analysis. Belmont, CA: Wadsworth; Maltz, M. D. 1984. Recidivism. New York: Academic Press; Schmidt, P., Witte, A. D. 1984. An Economic Analysis of Crime and Justice: Theory, Methods, and Applications. Orlando, FL: Academic Press; Schmidt, P., Witte, A. D. 1988. Predicting Recidivism Using Survival Methods. New York: Springer-Verlag; Tuma, N. B., Hannan, M. T., Groeneveld, L. P. 1979. Dynamic analysis of event histories. American Journal of Sociology 84:820-54.

[21] Barton, R. R., Turnbull, B. 1979. Evaluation of recidivism data: use of failure rate regression models. Evaluation Quarterly 3:629-41; Carr-Hill, G. A., Carr-Hill, R. A. 1972. Reconviction as a process. British Journal of Criminology 12:35-43; Harris, C. M., Moitra, S. 1978. Improved statistical techniques for the measurement of recidivism. Journal of Research in Crime

One of the foremost deficiencies of the binomial approach to prediction is that time-related information is entirely ignored: rearrest for a serious violent crime is treated as a "success" if it occurs at <u>any</u> time during the fixed follow-up period or a "failure" if it does not occur during the entire follow-up period. (This dichotomous representation of the behavioral outcome reflects, of course, the binomial aspect of time-unrelated prediction.) Consider the following example of two young adults in the 1958 birth cohort: Both subjects were arrested for aggravated assaults on their twenty-first birthdays, and both were later rearrested, but one was rearrested just one month later and the other in the sixth month. Surely it is reasonable to hypothesize that the first birth cohort subject posed a worse risk of being rapidly rearrested than the second birth cohort subject, if one accepts that the more rapid time until rearrest reflected some stable, underlying aspect of the first subject's behavior rather than a mere chance event. Traditional time-unrelated prediction techniques uniformly failed to exploit temporal information, which can often shed light on the origins of differences among subjects in their comparative risks of rearrest over time, even though such information is commonly available and easily retrieved.

Another disadvantage of the binomial approach to prediction is that study subjects must be observed for the entire follow-up period or over their entire lifetimes in order to determine unequivocably their rearrest statuses. Those subjects who cannot be observed for the entire follow-up period (e.g.,

---

<u>and Delinquency</u> 15:194-213; Harris, C. M., Kaylan, A. R., Maltz, M. D. 1981. Recent advances in the statistics of recidivism. In <u>Models in Quantitative Criminology</u>, ed. J. A. Fox, 61-80. New York: Academic Press; Maltz, M. D., McCleary, R. 1977. The mathematics of behavioral change. <u>Evaluation Quarterly</u> 1:421-38; Stollmack, S., Harris, C. M. 1974. Failure rate analysis applied to recidivism data. <u>Operations Research</u> 22:1192-1205.

death, residential relocation) or over their entire lifetimes (e.g., withdrawal from the study), cannot have their rearrest statuses unequivocably determined, yet this information must be known in order for the approach to be properly used. Uncertainty in this regard reflects the problem of the "censored" observation. It may be impossible, for example, for researchers to track an arrested birth cohort subject for the entire follow-up period, let us say for twelve months, either because of a residential move which may have occurred in the ninth month or because the study terminated in that same month. The birth cohort subject may not have been rearrested by the ninth month, and this can be established with certainty. However, after the residential move or study's termination, the subject might have been rearrested elsewhere or at a later time, in a jurisdiction or during a time period not covered by the study. But, and this is the nub of the quandary, the subject might just as well <u>not</u> have been rearrested in that jurisdiction or during that later time period. Neither rearrest nor the lack thereof can, therefore, be verified with certainty. Consequently, the subject's rearrest status over the full twelve months is clouded. Traditional prediction approaches might either have excluded a case like this from the analysis or have used some other strategy which discards or distorts information. The rearrest-free period prior to the residential move or study termination can, however, be classified unambiguously as an abstinent period (at least relative to official detection) and provides useful information which could be incorporated into analyses were the time-related failure analysis approach adopted rather than the time-unrelated binomial approach.

The upshot of this discussion should now be clear: the failure time approach explicitly and centrally incorporates time into analyses and, as a significant byproduct, includes subjects who have been observed for different lengths of time, making for a more powerful, more precise, and more versatile analysis of time-related rearrest risks. The binomial approach, however, falls short on each count, forcing the exclusion of either subjects or data, resulting in a distorted assessment of rearrest risks.

Key Aspects and Applications of Failure Time Analysis

The distribution of the times until rearrest for violent crimes can be statistically described in three equivalent ways: (1) the hazard function, (2) the survival function and its complement, the failure function, and (3) the probability density function. While these functions are mathematically equivalent and can be converted into one another, they represent different features of the rearrest-time distribution and, therefore, have different practical and conceptual implications. The first two functions were the more useful ones in striving to reach an acceptable level of prediction accuracy.

The hazard function is, with respect to the present study, the time-conditional risk (i.e., rate) of rearrest for a serious violent crime. Put somewhat differently, the hazard function is the probability that a birth cohort subject who has been arrested for a violent index crime will be rearrested for another one of these crimes within some specified time interval, given that that subject has not been rearrested for a violent index crime by the start of the specified time interval.[22] Consider for a moment the following example. Front-line js decision makers are confronted with a

[22] For this reason, the hazard rate is sometimes called the age- or time-specific failure (i.e., rearrest) rate. The hazard rate is a conditional rate and is based only on those persons who are still at risk of rearrest at the start of the age or time interval of interest.

birth cohort subject who has been arrested for a violent crime on his twenty-first birthday. Over the next three months, there are no further arrests for violent crimes. The probability that this subject will be rearrested for a violent crime during, let us say, the next (fourth) month, given that the subject has remained arrest-free through the third month, is the rearrest hazard rate that the subject sustains during that next month. One can view the hazard function as generating the survival and failure functions described next; the temporal pattern in rearrest risks, expressed by the hazard function, generates the overall rearrest-free and rearrest-punctuated periods following an arrest. The hazard function is defined in figure 2.2 and is accompanied by its conventional computational formula.

The survival function is, also with respect to the present study, the probability that a birth cohort subject who has been arrested for a serious violent crime will remain free from rearrest for another serious violent crime past some specified time point; that is to say, the subject will "survive" beyond that time point. Consider once again the above example. The survival function represents the probability that a birth cohort subject who has been arrested for a serious violent crime on his twenty-first birthday will not sustain a rearrest for a serious violent crime until sometime after a specified age, let us say, his twenty-sixth birthday. The survival function is defined in figure 2.2 and is accompanied by its conventional computational formula.

The failure function, which is the complement of the survival function (i.e., one minus the failure function), represents the probability that a birth cohort subject who has been arrested for a serious violent crime will be rearrested for another one of these crimes before or at some specified time point. The birth cohort subject has "failed" to remain arrest-free prior to the specified time cutpoint. With respect to the present study, the failure function can aptly be termed the rearrest function. Aspects of this function are intensively examined in later analyses.

The probability density function, or just density function, is defined in the present study as the probability that a birth cohort subject who has been arrested for a serious violent crime will be rearrested for another serious violent crime during a later time interval of interest. Consider (for just one last time) the above example. The density function represents the probability that a birth cohort subject who has been arrested on his twenty-first birthday for a serious violent crime will be rearrested for a serious violent crime during, let us say, the third month after that birthday; or during the fourth month, or during some other later (or perhaps earlier) month.[23] The probability density function is defined in figure 2.2 and is accompanied by its conventional computational formula.

The forthcoming analyses have focused on the hazard and failure functions because of their greater utility in framing clear js policy issues and goals. The hazard function can help js decision makers pinpoint those times at which persons arrested for serious violent crimes will experience heightened risks of rearrest. This information can be used to enhance the attainment of violent crime prevention and control objectives: for example, supervision and social support services can be intensified during peak risk

---

[23] In contrast to the hazard function, the density function represents the unconditional probability of rearrest during a specified time interval because it does not stipulate that the subject must be free from rearrest at the start of that time interval. Consequently, those subjects who have not remained arrest free prior to the start of the time interval are included in the base of the density-function computation.

periods to help shepherd persons through these periods without renewed criminal incident.

The rearrest function can help js decision makers plan when to initiate or terminate supervision and social support services. For example, it seems sensible to begin supervision or to initiate the delivery of intensive social support services promptly after arrest for those persons who possess the highest risk of very rapid rearrest and who, conversely, possess the lowest risk of remaining free from rearrest ("surviving") for a long period of time. It also seems sensible to relax if not completely end supervision or support services for those persons who have remained arrest-free long enough to suggest that they have a very low remaining risk of being rearrested in the future. The rearrest function highlights when these periods begin and end.

To illustrate the last point, suppose that a group of birth cohort subjects have been arrested on their twenty-first birthdays for serious violent crimes and that the rearrest function indicates that each of these subject has a .90 risk of being rearrested within three years, by his twenty-fourth birthday. (This probability is their accumulated failure rate.) Conversely stated, each subject who has remained arrest-free through age twenty-four has a .10 risk of being rearrested thereafter. (This probability is their survival rate.) One might reasonably argue that a .10 risk of rearrest after some cutoff time represents a sufficiently low risk to support the decision to discontinue js supervision and services beyond that time. If the risk of rearrest is low enough, why not disengage the js from that person after this critical time has been reached? [24] Although the rearrest risk is

---

[24] This is a variant of the "critical time" approach used in some failure time studies. The critical time approach is based on the idea of a formal shift in a person's rearrest risk from one underlying process (i.e.,

not the sole basis for making this decision--the social and personal consequences of the anticipated recurrent behavior also commonly enters into the picture (these are thevarious stakes which are involved)--this risk is certainly a key ingredient.[25]

Prediction is a common and central application of the failure time approach described above. The extent to which the hazard and failure functions can be put to use in making routine but quite sensitive js decisions about how to process violent criminals will be determined in part by the capacity of these functions accurately to predict those seriously violent criminals who will be at greatest risk of rearrest at certain times. Prediction accuracy and generality (i.e., validity) will, therefore, be prominent themes in the following analyses.

Why Use Parametric Failure Time Models of Rearrest?

There are two approaches one can take to statistically describing the distribution of rearrest times: nonparametric and parametric. This study stressed the second approach.

The nonparametric approach attempts to represent--"match"--the patterns appearing in the observed (manifest) rearrest times, but the approach does not

---

distribution), generating a high risk of rearrest, to another underlying process, generating a low risk of rearrest, including absolute rehabilitation in which the person shifts to a process generating absolutely no risk of rearrest. In this study, however, we view matters less formally. When a person reaches a specified low level of rearrest risk, this might form a reasonable basis for ending or reducing js involvement with that person. One need not view the low level of risk as a formal shift in distributions. For details about "critical time" analysis, see Maltz, M. D., 1984. Recidivism. Orlando, FL: Academic Press, Inc.

[25] See Gottfredson, S. D., Gottfredson, D. M. 1988. Stakes and risks in the prediction of violent behavior. Violence and Victims 3,4:247-62.

assert that the underlying theoretical distribution characterizing the observed distribution of rearrest times has a particular shape (i.e., curvilinear form). For this reason, this approach is described as "distribution-free." Because no specific underlying theoretical distribution is asserted, no distributional parameters need to be estimated, which is the basis for referring to the approach as "nonparametric."[26]

The parametric failure time approach, however, does assert that the underlying theoretical distribution characterizing the observed distribution of rearrest times has a specific shape. By making this assertion, the parametric approach argues that the distribution of observed rearrest times has been generated by an underlying behavioral process and, furthermore, that this process can be described by a mathematical equation representing a specific failure time distribution. This failure time distribution is represented by a structure of estimated parameters and coefficients. The overall strength and utility of the parametric failure time approach can be realized only to the extent that the "correct" underlying distribution is selected. If the essential features of the observed rearrest times is not reflected by the selected underlying distribution, that distribution can misdirect both theory-building and js policies and practices, a potential pitfall, one might add, of parametric modeling in general, not just of failure time modeling. If, however, the essential features of the rearrest times are

[26] For discussions of these methods, see Berkson, J., Gage, R. R. 1950. Calculation of survival rates for cancer. Proceedings of Staff Meetings, Mayo Clinic 25:252; Cutler, S. J., Ederer, F. 1958. Maximum utilization of the life table method in analyzing survival. Journal of Chronic Diseases 8:699-712; Gehan, E. A. 1969. Estimating survival functions from the life-table. Journal of Chronic Diseases 21:629-44; Kaplan, E. L., Meier, P. 1958. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53:457-81.

indeed reflected by the selected distribution, that fact can yield significant benefits.

One benefit of the parametric failure time approach is that it economically represents the underlying behavioral process which might be governing the observed rearrest times rather than simply matching on an _ad hoc_ basis the observed rearrest times themselves. The parametric approach involves first asserting that the observed distribution of rearrest times possesses a specific curvilinear shape and then estimating the parameters and coefficients defining that curvilinear shape. The two steps--curve _specification_ and curve _estimation_--result in a parametric statistical "model" of the underlying behavioral process which generated the observed rearrest times, and it functions to smoothe the frequently erratic shape of the distribution of these observed times. The resultant statistical model, based on finite observed data, can then be used to estimate the conditional probability (risk) of rearrest within a specified time period (the hazard rate), the probability that rearrest will not occur before a specified time period has elapsed (i.e., the survival rate), or the probability that rearrest will indeed occur before a specified time period has elapsed (i.e., the failure rate). Furthermore, one can make these kinds of estimations and, based on them, predictions for time periods extending _beyond the finite range of the observed rearrest times_.

This extendable aspect of the parametric failure time approach is extremely useful because it helps to loosen the fetters of limited, finite data. For example, in the present study, arrest records were unavailable after the subjects' twenty-seventh birthdays. The parametric failure time approach, however, would enable the computation of hazard rates and failure rates beyond this upper age limit. Furthermore, this approach permits the analytical results to be applied to other, diverse populations. For instance, a parametric model of rearrest which has been estimated using subjects in one birth cohort can be employed in making predictions about the rearrest behavior of subjects in another birth cohort. Importantly, the validity of the failure time models developed using one birth cohort can be partly evaluated, by assessing the predictive accuracy of the models, when used for making predictions about rearrests in another birth cohort.

For reasons which should by now be clear, the present study mainly adopted the parametric failure time approach to studying violence. First, this approach resulted in an economical representation of the rearrest times. Second, this approach permitted the generalization of results beyond the limited range of the observed rearrest times. Third, because the parametric approach enables and compels consideration of what variables or processes influence rearrest risks at different times, the approach promoted a focused discussion of the dynamic aspects of public policy strategies and related theoretical concerns, some of which has already appeared in these pages.

While the parametric failure time approach formed the core of the study, the nonparametric approach was also used, but in a limited way. Nonparametric computations of, for example, the hazard rate provided a useful benchmark against which to compare the accuracy of the corresponding parametric estimates. Also, the partially parametric, Cox proportional hazards regression model was used to help gauge the plausibility of the fully parametric regression models because it yields fairly consistent (i.e., robust) results across distributional forms.

<u>Selecting the Parametric Failure Time Distributions</u>

Parametric failure time distributions can sharply differ in their shapes, and this is probably most clearly seen with respect to the (age- or time-) conditional risk of rearrest expressed by the hazard function. Commencing with an arrest for a serious violent crime, the hazard function can, thereafter, continuously increase, decrease, or remain constant; furthermore, these increases or decreases can be either linear or curvilinear. Alternatively, the hazard function can first increase and then decrease; it can behave in just the reverse fashion; or it can show even more complex shapes. In fact, there are as many possible shapes to the hazard function as there are parametric failure time distributions.

It is difficult to think of a single pattern in the hazard function that might apply across the spectrum of violent crimes and criminals. Many patterns seem plausible, depending upon one's theory or speculation about the commencement, continuation, and cessation of such behavior. Consider the cessation issue. Assume that an arrest deters, over the short run, the commission of another violent crime after an offender has returned to civil society. The arrest's deterrent impact might, then, result in a low initial rearrest risk, followed by a higher rearrest risk as time passed and the deterrent effect subsided. Now alternatively assume that an arrest rehabilitates, over the long run, the violent criminal. The arrest may facilitate, for any number of reasons (e.g., moral awakening, personal insight) an increased capacity over time to refrain from another violent crime. The arrest's rehabilitative impact might result, then, in a high initial rearrest risk, followed by a lower rearrest risk.

One can easily think of other plausible patterns. Consider the implication of the following aspects of robberies and assaults: their instrumental versus expressive motivations. Robberies are often viewed as mainly instrumental crimes, as motivated by the strategic acquisition of valuables, such as money. In this respect, robberies are akin to other economically motivated activities. If instrumentality is, indeed, the dominant aspect of robberies, one might reasonably assert that once a person has been arrested for a robbery, that person's need to acquire valuables rekindles, increasing over time; monetary depletion motivates monetary replenishment. This increasing need to acquire money is transformed into a progressively increasing risk of robbery. A hazard function which increased over time would have to be invoked to match this pattern in rearrest risks.

Assaults, on the other hand, are often viewed as mainly expressive crimes, for example, as representing emotional responses to certain kinds of interpersonal provocations or triggers. These provocations commonly recur on a regular basis (e.g., spousal arguments resulting from the frequent intoxication of one partner; disputes between barroom buddies each payday). The risk of rearrest for assaults resulting from these provocations might be fairly level over time, given the great frequency and regularity of these provocations. A constant hazard function would be needed to match this pattern.

We see that one can make the case that some violent criminals remain continuously on the brink of rearrest, that others must first build up steam before approaching this brink, and still others, from the word go, begin to lose steam and move away from the brink. (Other, more complex, conjectures are also conceivable.) In light of these possibilities, and of precious little systematic theory specifically about violent crimes to which one can turn for guidance, it is easier to posit an argument in favor of using a particular parametric distribution than of not using it. Because statistical software now exists which permits the examination of many different distributions, there is no need to skimp in this regard. We adopted, therefore, an eclectic approach to reflecting the diversity of violent phenomena, selecting several parametric distributions for examination. A few simple criteria were used to guide the selection of these distributions beyond the most basic criterion that the distribution must be nonnegative (because rearrest times can only be positive): precedence, versatility, economy, interpretative diversity, and interpretative clarity.

First, failure time distributions were selected which had proven useful, practically and theoretically, in related prior studies. We hoped to capitalize on prior research and, thereby, to build upon past successes, for example, by including some nonmonotonic distributions (i.e., represented by a curvilinear hazard function) and some distributions which are skewed to the right (some research, for example on parole failure, has found that recidivism seems disproportionately to occur early, to increase quickly, and then to taper off quickly). However, virtually no prior research has focused on arrests for serious violent crimes, nor used a sequential-prediction framework for examining the related rearrest dynamics. (This scarcity is more fully discussed later.) Thus, research bearing directly on the present study is quite limited and provides only the broadest counsel about potentially pertinent failure time distributions.

Second, the hazard functions of the selected failure time distributions had to provide a versatile set of shapes. By ensuring wide coverage in this respect, we promoted a broad search for the most appropriate hazard functions relating to rearrest for violent crimes rather than foreclosing on this search by prematurely limiting the shapes of candidate distributions.

Third, the selected failure time distributions had to permit the economical representation of the times until rearrest for violent crimes. Whenever possible, simpler distributions, those with fewer parameters to estimate, were selected to act as foils to their more complex and less parsimonious general ("parent") distributions.

Fourth, failure time distributions were chosen to reflect alternative behavioral interpretations of the dynamics underlying rearrests for serious violent crimes. This is, of course, an issue of construct validity and is expressed in the following question: Does a selected parametric model plausibly reflect some of the most salient and central behavioral dynamics governing rearrests for violent crimes? For example, one class of failure time distributions formally implies that all persons arrested for violent crimes will eventually be rearrested for such crimes, if given enough time (i.e., a homogeneous, or <u>unitary-population</u>, parametric model). Another class of failure time distributions formally implies that some persons arrested for serious violent crimes will eventually be rearrested for these same crimes, if observed long enough, but that other persons will not be rearrested for these crimes, no matter how long they might be observed, because they have permanently ended their involvement in violent behavior (i.e., a heterogeneous-, incomplete-, or <u>split-population</u>, parametric model, sometimes also called the desistance model).

Unitary-population failure time distributions assume that all persons are behaviorally "susceptible" to rearrest because they engage in criminally violent behavior which places them at risk of rearrest for that behavior. Split-population distributions, on the other hand, assume that some persons are behaviorally "immune" to rearrest for violent crimes, whereas other persons are "susceptible" to rearrest for these same crimes; the immune group either will not engage in violent behavior or, if they do engage in such behavior, will not be arrested for that behavior, although how this might happen is unclear. To begin to explore the empirical merits of these quite different behavioral interpretations, whenever possible, both the unitary- and split-population parametric distributions were employed in the present study. As we assess these two types of distributions, we will bear in mind, however, the proper caution that a distribution's "realism" (i.e., construct validity) may not directly correspond to its "utility."[27] The more complicated split-population parametric distributions, ostensibly the more realistic ones, may sometimes offer little, if any, bonus in either statistically describing the observed failure times (the model's "fit") or in understanding the underlying dynamics which generated these times (the model's interpretation). Realism may be possible to achieve only when the data requirements of models using split-population parametric distributions can be met: for instance, the rearrest rate must be fairly high and the observation period must be sufficiently long.[28] While the first of these broad rules is often met by the birth cohort data, the second rule may not be. The eight-year time spans in the juvenile and young adult periods may not be adequate to discern firmly the relative merits of the unitary- and split-population distributions.

[27] Rhodes, W. 1990. The criminal career: estimates of the duration and frequency of crime commission. Journal of Quantitative Criminology 5,1:3-32.

[28] Rhodes, W. 1990. The criminal career: estimates of the duration and frequency of crime commission. Journal of Quantitative Criminology 5,1:30-31.

Fifth, failure time distributions were selected whose parameters had the most straightforward behavioral and conceptual interpretations. For example, the more attractive failure time distributions were those whose parameters clearly described the curvilinearity of the hazard function (i.e., the shape parameter) or the division of the at-risk population into distinct behavioral subgroups, such as an "immune" group and a "susceptible" group (i.e., the splitting parameter).

Motivated by the above considerations, a wide array of unitary- and split-population parametric distributions were incorporated into this study: the Weibull, lognormal, and loglogistic; the extreme-value and split population versions of each of these; the mixed exponential; and the Gompertz. For more detailed treatments of the technical aspects of these distributions and their functions, one can refer to basic texts on failure time distributions,[29] the literature on mathematical and statistical applications in criminology and criminal justice,[30] and documentation accompanying the major statistical computing packages.[31]

---

[29] Allison, P. D. 1985. Event History Analysis. Beverly Hills: Sage; Cox, D. R., Lewis, P. A. W. 1966. The Statistical Analysis of Series of Events. London: Methuen; Cox, D. R., Oakes, D. Analysis of Survival Data. New York: Methuen; Kalbfleisch, J. D., Prentice, R. L. 1980. The Statistical Analysis of Failure Time Data. New York: Wiley; Lawless, J. F. Statistical Models and Methods for Lifetime Data. New York: Wiley; Lee, E. 1980. Statistical Methods for Survival Data Analysis. Belmont, CA: Wadsworth.

[30] Maltz, M. D. 1984. Recidivism. New York: Academic Press; Schmidt, P., Witte, A. D. 1984. An Economic Analysis of Crime and Justice: Theory, Methods, and Applications. Orlando, FL: Academic Press; Schmidt, P., Witte, A. D. 1988. Predicting Recidivism Using Survival Models. New York: Springer-Verlag.

[31] Maltz, M. D. 1991. Survival Fitting and Analysis Software for Industrial, Biomedical, Correctional, and Social Science Applications. University of Illinois at Chicago Circle; Steinberg, D., Colla, P. 1988. SURVIVAL: A Supplementary Module for SYSTAT. Evanston, IL: SYSTAT, Inc.

Determining the "Best" Failure Time Distribution

A thorny problem often faced when using several parametric failure time distributions is how to decide which one best matches the observed array of rearrest times. This decision can be guided by both formal and practical considerations. The practical approach was emphasized because because we are primarily interested in assessing the utility of a distribution when applied to predictive decision making. We turn first to the formal considerations.

The relative appropriateness of rival distributions was formally evaluated by comparing their loglikelihoods. That underlying distribution with the highest loglikelihood (i.e., the least negative) was deemed the most appropriate distribution in the sense that it was judged to have been the most likely one to have generated and, thereby, to have "explained" the observed array of rearrest times. This strategy did not assert, however, that the highest loglikelihood value was discernibly highest in the technical statistical sense, based on formal hypothesis testing, but only that the loglikelihood was highest in the absolute numerical sense. In theory, such hypothesis testing can only be performed on distributions which are members of the same distributional family. Specific members of a family can be contrasted sequentially to the most general member of that family to decide which specific distribution, if any, is statistically the best. Unfortunately, this strategy was not feasible in the present context. Many of the distributions selected for study did not belong to the same family of distributions, and those which did so could not easily be contrasted due to current limitations in statistical computer software. Matters became even more intractable when risk variables were introduced into the analysis: there were many "unnested" regression models, making it impossible systematically to compare them.

While the loglikelihood criterion was not exclusively invoked to select the "best" failure time distribution, it nevertheless did provide a useful starting point in making the selection. The first step in this process was formally to identify that failure time distribution possessing the highest loglikelihood; the second step was to evaluate the identified distribution within the practical framework of predictive decision making. In the present study, then, one of the acid tests of a distribution's appropriateness was, quite simply and starkly, its level of predictive accuracy and whether that level might have some practical utility as an aid in handling persons arrested for serious violent crimes. This central principle for establishing a distribution's credentials was as it should be given the great emphasis this study placed on applied utility. We wanted to steer clear of spending alot of time splitting technical statistical hairs formally targeting the "best" distribution if even the best one was unsatisfactory from the standpoint of practical decision making.

The Problem of Censored Cases

As was briefly noted earlier, the censored observation is one of the thornier but, nowadays, quite tractable problems faced by researchers studying time-related behavioral outcomes like the times between arrests. If a birth cohort subject was not rearrested, let us say for a second serious violent crime, by the close of the observation period, for example, by his eighteenth birthday, one can confidently say that that subject was not rearrested prior to that time; but one cannot confidently say whether a rearrest occurred after that time barrier was reached. However, the potential for rearrest clearly does not cease to exist beyond the bounds of the study's observation period. The study's ending date, therefore, prematurely terminates an ongoing arrest career which, in fact, continues beyond that date; hence, the ending date is a methodological imposition, not a behavioral event. When characterizing rearrest risks, one tries to describe this process as it continues over the entire time period during which persons are at risk to be rearrested--a lifetime--rather than up to some time cutpoint which has been artificially imposed by constraints on the study design.

In the present study, every birth cohort subject who was arrested for a serious violent crime could potentially have been censored at any point in the arrest sequence, either at age 18, if the focus was on juvenile arrests, or at age 26, if the focus was on young adult arrests.[32] How does one deal with the potentially ambiguous rearrest statuses of those subjects who were not rearrested by the time the age cutpoint was reached?

Several approaches to dealing with the censoring problem have been proposed.[33] However, some of these strategies are unattractive because they unavoidably weaken and distort findings. For instance, one strategy might be to exclude censored rearrest times from the study and only include those times reflecting actual rearrests for violent crimes. However, censoring may be substantial, resulting in too few actual rearrest times to support reliable

---

[32] In fact, every birth cohort subject who was arrested was censored at least once in this study, either at age 18 or at age 26. This is so because every subject, by definition, had a final arrest in the observed arrest sequence which was then followed by one of these two censoring points.

[33] For example, see the discussion by Tuma, N. b., Hannan, M. T. 1978. Approaches to the censoring problem in analysis of event histories. In Sociological Methodology, 1978, ed. K. F. Schuessler, 209-40. San Francisco: Jossey-Bass.

analyses. The exclusion of data creates another problem, estimation bias. Note that uncensored rearrest times are those times which occurred prior to the intervention of the cutoff age. By excluding the censored rearrest times, which represent the potentially longer rearrest times, this approach underestimates the expected times until rearrest and, by extension, overestimates the hazard rate. If one were unwittingly to apply the products of such biased estimation to predictive decision making, one would incorrectly predict shorter time intervals between arrests than would actually occur. The extent of this bias would, of course, depend upon the underlying process which generated the observed rearrest times.

A second, also unattractive, approach to handling censored rearrest times is to act as if these censored times signified actual rearrests which occurred precisely at the time the censoring barrier intervened (in the present study, at the cohort subjects' 18th and 27th birthdays). Unfortunately, this approach also induces estimation bias because it creates artificial rearrest times which are less than those which would actually have occurred. As with the above approach to handling censoring, this one also underestimates the time until rearrest and, thus, also overestimates the hazard rate.

A third--and the preferred--approach is to "...employ a method of estimation that adjusts for censoring under the assumption that the same stochastic [probabilistic] model applies to all cases, whether or not observations of them are censored."[34] The advantage of these adjustments is that the full complement of time-related information which is available about

---

[34] Tuma, N. B., Hannan, M. T. 1979. Approaches to the censoring problem in analysis of event histories. In Sociological Methodology, 1979, ed., K. F. Schuessler, 209-40. San Francisco, CA: Jossey-Bass, p. 213.

a birth cohort subject's history of arrests for serious violent crimes, spanning from the immediate arrest until the censoring cutoff time has been reached, is efficiently preserved and utilized rather than discarded: once the subject can no longer be observed due to censoring, that subject is removed from the overall pool of subjects who are still at risk of being rearrested but without assuming that that subject will later be rearrested. This procedure permits one to employ of information on rearrest times about <u>both</u> censored and uncensored subjects because the information about both types of subjects is assumed to describe the same underlying theoretical distribution extending up to and beyond the censoring barrier.

<u>Incorporating Risk Variables into the Failure Time Models</u>

Whenever we used a particular parametric failure time distribution, we <u>initially</u> proceeded with the analysis <u>as if</u> the birth cohort subjects formed a <u>homogenous</u> group because we assumed that the observed rearrest times of all the subjects had been generated by a <u>common</u> underlying behavioral process reflected by that particular distribution. For example, when we used the Weibull distribution, we assumed that the times until rearrest of <u>all</u> birth cohort subjects were governed by that distribution. This was the basis for fitting the most appropriate Weibull distribution to the <u>entire</u> array of rearrest times. In addition to fitting a parametric distribution to the rearrest times and, thereby, describing (i.e., modeling) the overall (i.e., marginal) probabilistic <u>shape</u> of these times, the influence of an <u>individual's</u> exposure to risk variables was also examined in order to trace the comparative (i.e., conditional) magnitudes of rearrest risks of each birth cohort subject.[35] Essentially, we asked the following question: Which criminal history and personal characteristics (these are represented by the risk variables) of individuals who have been arrested for violent crimes are related to a high probability of rapid rearrest? For example, Does it matter what type of violent crime the individual was arrested for? Whether a firearm was present? When the individual was first arrested for a violent crime? When the individual was last arrested for such a crime? Whether the individual had ever been confined in a secure facility? These and similar questions formed the bulwark of the study, and they were the basis for trying to strengthen predictive decision making.

In selecting risk variables for the analysis, we were interested only in those ones which were predictive of rearrest for a violent crime at the time of the instant arrest because it is at that time that front-line js decision makers must make their decisions about how to handle the arrested person. This proviso places a potentially daunting requirement on the chosen variables; they must characterize the person at the time of arrest and be predictive of future rearrest. For this reason, the present study did not use time-varying risk variables which might have characterized the personal features and criminal history of the birth cohort subject between the time of the instant arrest and the time of the rearrest. In short, we did not attempt to explain the dynamics of rearrest, extending from the time of the instant

[35] The analysis of failure times which omits risk variables involves the analysis of the marginal distribution of these times, whereas the analysis of failure times which includes risk variables involves the analysis of the conditional distribution of these times. See Schmidt, P., Witte, A. D. March 1987. Some thoughts on how and when to predict in criminal justice settings, working paper. Wellesley College, Wellesley, MA.: Department of Economics, p. 17.

arrest through the time of the rearrest; rather, we did attempt to predict rearrest by identifying risk variables which were related to rearrest at the time of the instant arrest.

Identifying influential risk variables is important for two reasons, one that is obvious and one that is not. First, these risk variables increase the reliability of results by reducing the variance of the prediction estimate for each individual, consequently strengthening the prospects for more accurate individual and aggregate prediction.[36] (This is the obvious reason for trying to identify influential risk variables.) Second, these risk variables enhance the practical utility of the parametric models when applied to other, nonrandom samples of subjects. Risk variables enable one partially to correct for differences in the personal and criminal history characteristics between the group initially used to develop the prediction model (the construction sample) and the group on which the prediction model is subsequently used (the validation or application sample).[37] (This is the less obvious reason for trying to identify influential risk variables.) As one might expect, the usefulness of a prediction tool will be broadened to the extent that the characteristics of subjects used in the construction sample reflect those of the wider groups on which one would like to use the results. This consideration pertains to the familiar issue, discussed earlier, of the

---

[36] Schmidt, P., Witte, A. D. March 1987. Some thoughts on how and when to predict in criminal justice settings, working paper. Wellesley College, Wellesley, MA.: Department of Economics, p. 19; Schmidt, P., Witte, A. D. 1988. Predicting Recidivism Using Survival Models. New York: Springer-Verlag.

[37] Schmidt, P., Witte, A. D. March 1987. Some thoughts on how and when to predict in criminal justice settings. Working Paper, Department of Economics, Wellesley College, Wellesley, MA., pp. 19-20; Schmidt, P., Witte, A. D. 1988. Predicting Recidivism Using Survival Models. New York: Springer-Verlag.

external representativeness of the construction sample, which can be initially bolstered through wise sampling and subsequently (partly) corrected for, when deficient, through the use of relevant risk variables.

To accomodate risk variables and, by so doing, to take into account personal and official information about the birth cohort subjects which might be related to more compressed times between their arrests for serious violent crimes, multivariate failure time (regression) models were adopted. These multivariate models permitted introducing into the analyses the spectrum of ethically and legally permissible and legally impermissible risk variables detailed earlier. In view of the regression framework, one should not be surprised to learn that the risk variables were hypothesized to affect the average, or common, underlying hazard rate at a particular time, displacing this rate upward or downward according to a weighted combination of the risk variables characterizing a particular birth cohort subject. For example, when applying a Weibull distribution with a decreasing hazard function, we assumed that all arrested birth cohort subjects exhibited hazard rates which followed parallel decreasing paths (the shape aspect), but the hazard function of one subject might always be well above that of another subject (let us say, twice as much) because certain variables characterizing that subject amplified that subject's rearrest risk (the magnitude aspect).

By incorporating risk variables into the analysis, each birth cohort subject was, in effect, characterized by his own hazard function. In this way, individual differences among birth cohort subjects were acknowledged which might be related to the risk and timing of their rearrests for violent crimes. The multivariate failure time models functioned, then, as predictive prisms, first ranking differential risks of rearrest across persons and then, for each person, ranking differential risks of rearrest across time. With this kind of risk-ranking information at hand, js decision makers might be in a better position to make more effective choices about individual dispositions. By individualizing the rearrest risks, the multivariate models had the additional salutary effect of promoting the more reliable transference of these models to other, nonrandom samples.

The Generality of the Failure Time Models: Split Samples and Dual Cohorts

The study findings will stand a far better chance of being fairly considered for adoption by front-line decision makers in diverse jurisdictions if these decision makers are presented with some hard evidence of the capacity, or the lack thereof, of the multivariate failure time models to predict accurately rearrests for serious violent crimes in populations which are distant in time and place from the original study group. Documented generality of the study findings is the only way to provide some cautious reassurance (if warranted) that these findings are not unduely limited in their application to the original study group. This is one reason why the validation of findings was so strongly emphasized by the present study.

One often employed and effective procedure for initially gauging validity is judiciously to use the original study population itself. The population is randomly divided into two subsamples (or, even more subsamples, depending upon research exigencies): a construction sample and a validation sample. The "best" multivariate failure time model is developed using the construction sample and is tested for its adequacy on the validation sample. This study was fortunate to have had a sufficient number of subjects to adopt this split-sample design.[38] Multivariate failure time models were first estimated using the larger construction group drawn from the 1958 birth cohort. The validity of these models was then assessed by using them to predict rearrests in two application groups: (1) the validation sample of the 1958 birth cohort and (2) the total 1945 birth cohort (for the juvenile period) and the follow-up sample (for the adult period). In each application, both individual and aggregate predictive accuracy were assessed.

---

[38] The 1958 birth cohort was split into a construction group (70 percent) and a validation group (30 percent). This allocation was guided by two considerations. First, we wanted to ensure that the construction group had the majority of subjects in order to strengthen the reliability of the initially estimated models. Second, it was desireable to develop failure time models across the more advanced points in the arrest sequence and, furthermore, to be able to assess the validity of these models. The 70-30 split made these extended analyses possible.

Once the birth cohort subjects had been randomly assigned to the two groups, we wanted to explore whether the assignment procedure had inadvertently produced systematic biases, making these groups incomparable with respect to some key predictive characteristics. Whatever differences might be present between the two groups should be due to the sampling variation inherent in the random assignment procedure and not due to any biasing peculiarities of the groups. To explore this issue, we ran a series of MANOVAs using three three-variable clusters to gauge group similarities: (1) the <u>total number</u> of arrests, violent index-crime arrests, and property index-crime arrests, (2) the <u>total seriousness</u> of all arrests, violent index-crime arrests, and property index-crime arrests, and (3) the <u>age at onset</u> of arrest, violent index-crime arrest, and property index-crime arrest. The MANOVAs failed to detect significant differences between the construction and validation groups.

The split-sample strategy is a <u>conservative</u> approach to assessing the validity of prediction models. On the one hand, it is unlikely that one will find greater model stability or predictive accuracy than that which is observed across the split-sample groups. On the other hand, although the split-sample strategy reduces the chances of invalidating models, some model invalidity will nevertheless be detected. The observed invalidity is a baseline against which one can compare other predictive applications. One is likely only to do <u>worse</u> when making these other applications, and if the split-sample results are dismal, matters will only get worse from that point on.

These procedures, applied both within and across the two birth cohorts, helped reduce the potential danger of overstating the extent to which predictive relationships which obtained in the construction sample also obtained in other groups. Overall, these procedures promoted a more realistic, conservative expectation about the level of analytical validity and predictive accuracy of the developed multivariate failure time models.

WHAT DOES THE PREVIOUS RESEARCH SAY?

The discussion which follows of prior research studies is brief--unfortunately, much too brief--in view of the serious types of crimes considered here and the powerful statistical tools now available to study these crimes. Quite simply, there does not appear to be a single study targeting seriously violent criminals which has used failure time techniques to track the probability and timing of their successive arrests for serious violent crimes. As one obvious consequence of this vacuum, there are also no studies <u>predicting</u> the probability and timing of these violent incidents, a fact which is surprising and frustrating, from both a scientific and public policy standpoint. These <u>are</u>, after all, arrests for some of the most serious crimes of some of the most serious criminals. This vacuum is all the more perplexing, and unsettling, considering that arrests for many types of serious violent crimes are not uncommon (robberies and aggravated assaults numerically dominate the picture) and are for behaviors which have, because of their great harm, spurred widespread and intense public fear and repugnance.

Despite the significance of these serious crimes, the impressive and often contentious research literature employing failure time techniques in criminology and criminal justice, roughly twenty studies in all, clustering in the last ten years, has largely either ignored or not focused on the specific topic of violent crimes.[39] Rather, this literature has overwhelmingly

[39] Barton R. R., Turnbull, B. W. 1979. Evaluation of recidivism data: use of failure rate regression models. Evaluation Quarterly 3,4:629-42; Barton, R. R., Turnbull, B. W. 1981. A failure rate regression model for the study of recidivism. In Models in Quantitative Criminology, ed. J. A. Fox, 81-101. New York: Academic Press; Bloom, H. S. 1979. Evaluating human service and correctional programs by modeling the timing of recidivism. Sociological Methods and Research 8,2:179-208; Carr-Hill, G. A., Carr-Hill, R. A. 1972. Reconviction as a process. British Journal of Criminology 12:35-43; Greenberg, D. F. 1978. Recidivism as radioactive decay. Journal of Research in Crime and Delinquency 15:124-25; Harris, C. M., Kaylan, A. R., Maltz, M. D. 1981. Recent advances in the statistics of recidivism measurement. In Models in Quantitative Criminology, ed. J. A . Fox, 61-80. New York: Academic Press; Harris, C. M., Moitra, S. D. 1978. Improved statistical techniques for the measurement of recidivism. Journal of Research in Crime and Delinquency 15:194-213; Holden, R. T. 1985. Failure time models for thinned crime commission data. Sociological Methods and Research 14,1:3-30; Lloyd, M. R., Joe, G. W. Recidivism comparisons across groups: methods of estimation and tests of significance for recidivism rates and asymptotes. Evaluation Quarterly 3,1:105-17; Maltz, M. D. 1984. Recidivism. Orlando, FL: Academic Press, Inc.; Maltz, M. D., McCleary, R. 1977. The mathematics of behavioral change: recidivism and construct validity. Evaluation Quarterly 1,3:421-38; Maltz, M. D., McCleary, R. 1978. Rejoinder on "stability of the parameter estimates in the split population exponential distribution". Evaluation Quarterly 2,4:650-55; Maltz, M. D., McCleary, R., Pollock, S. P. 1979. Recidivism and likelihood functions: a reply to Stollmack. Evaluation Quarterly 3,1:124-31; Partanen, J. 1969. On waiting time distributions. Acta Sociologica 12:132-43; Rhodes, W. 1986. A survival model with dependent competing events and right-hand censoring: probation and parole as an illustration. Journal of Quantitative Criminology 2,2:113-37; Rhodes, W. 1989. The criminal career: estimates of the duration and frequency of crime commission. Journal of Quantitative Criminology 5,1:3-32; Schmidt, P., Witte, A. D. 1980. Evaluating correctional programs: models of criminal recidivism and an illustration of their use. Evaluation Review 4,5:585-600; Schmidt, P., Witte, A. D. 1988 Predicting Recidivism Using Survival Models. New York: Springer-Verlag; Schmidt, P., Witte, A. D. 1989. Predicting criminal recidivism using "split population" survival time models. Journal of Econometrics 40:141-59; Stein, W. E., Lloyd, M. R. 1981. The Maltz-McCleary model of recidivism: a reexamination. Evaluation Review 5,1:132-44; Stollmack, S. 1979. Comments on "the mathematics of behavioral change". Evaluation Quarterly 3,1:118-23; Stollmack, S., Harris, C. M. 1974. Failure-rate analysis applied to recidivism data. Operations Research 23:1192-1205; Visher, C. A., Linster, R. L. 1990. A survival model of pretrial failure. Journal of Quantitative Criminology 6,2:153-84; Visher, C. A., Lattimore, P. K., Linster, R. L. 1991. Predicting the recidivism of serious youthful offenders using survival models. Criminology 29,3:329-66; Witte, A. D., Schmidt, P. 1977. An analysis of recidivism, using the truncated lognormal distribution. Applied Statistics 26,3:302-11.

concentrated on technical, methodological, and practical issues, not on behavioral ones like criminal violence: the technical merits of different types of statistical distributions, properties of estimators, and methods of significance testing; rival behavioral interpretations (i.e., construct validity) of different statistical models (i.e., unitary- vs. split-population); and the practical uses of these techniques (e.g., evaluating the benefits of competing criminal justice intervention programs). Maltz (1984), Schmidt and Witte (1988), and Rhodes (1989) have cogently and critically reviewed the strides made in these technical areas. Because these particular issues have historically drawn the most attention, there is virtually no research using failure time methods to consult specifically about criminal violence.

Several things seem to account for the thin research activity in this area. First, large samples are needed to net enough arrests for violent crimes to sustain reliable statistical analyses. Netting sufficiently large samples is a daunting enterprise, because both costly and time consuming, deterring researchers who might otherwise consider examining this topic. Second, and extending the previous point, considerable information is usually needed about the behavioral and cjs components of the arrest sequence, which is also both costly and time consuming to gather, doubly deterring potential researchers. Third, until quite recently, perhaps the last decade or so, only a handful of researchers studying issues relating to crime and the cjs were acquainted with more than the details of failure time statistical techniques.

Failure time studies on crime and the cjs passed through three broad methodological and conceptual phases (if one can even talk about "phases" when discussing just twenty studies): (1) those studies which used mainly unitary-population distributions but no risk variables, (2) those studies which used both unitary- and split-population distributions but no risk variables, and (3) those studies which used both unitary- and split-population distributions and risk variables. While we do not review in detail these study phases nor results, several of their more salient aspects are worth noting, in particular, analytical omissions relevant to our present concerns.

Some of the earliest failure time studies, for example, by Carr-Hill and Carr-Hill (1972) and Stollmack and Harris (1974), examined the length of time until criminal reinvolvement, measured as a reconviction for a new crime after release form prison (Carr-Hill and Carr-Hill) and as the violation of parole conditions (Stollmack and Harris). Neither study, nor others conducted around that time, however, separately examined violent criminals, the time between arrests for violent crimes, the comparative timing of successive arrests for violent crimes in the arrest sequence, or the influence of risk variables on the probability and timing of these arrests.

Even during this early phase, controversy sparked over which type of parametric distribution adequately reflected the behavioral dynamics underlying the timing of criminal reinvolvement (i.e., construct validity). For example, Stollmack and Harris employed a unitary-population model (specifically, the exponential model), but, not long afterward, Maltz and McCleary (1977) questioned whether a split-population model might not be more appropriate, in terms of both its technical adequacy (the statistical "fit") and its conceptual cogency ("construct validity"), because some arrested criminals might reasonably be expected <u>never</u> to repeat criminal acts, an outcome which is formally built into the split-population model. In their own study, however, and following in the steps of all previous researchers, Maltz and McCleary did not specifically examine violent criminals, their arrests for violent crimes, nor, by implication, potentially influential risk variables.

Few failure time studies have examined the effects of risk variables on the probability of reinvolvement in crime or with the cjs, or on the timing of these reinvolvements. The only studies which included risk variables into analyses were those by Barton and Turnbull (1981), Schmidt and Witte (1979; 1988), Rhodes (1989), and Visher and Linster (1990).[40] Furthermore, with precious few exceptions, these studies have used different sets of risk variables, frustrating comparisons of their effects across different techniques and samples. (Indeed, of the more than 25 distinct risk variables considered by these studies, only "age" was included in all studies. No other risk variable was included in more than two studies, and there was just a handful of these variables.) The generality of multivariate effects across time periods, locations, samples, and statistical techniques cannot,

---

[40] Barton and Turnbull (1981) studied the effects of the following risk variables on the timing of rearrest after release on parole: institutional placement, previous major offense, age at release, drug use, and monthly income. Schmidt and Witte (1988) studied the effects of the following risk variables separately on the probability and the timing of reimprisonment after release on parole: time served in prison, age at release, number of prior imprisonments, number of prison rule violations during the sample imprisonment, number of years of formal schooling, race, gender, serious alcohol problem, prior use of hard drugs, marriage status, parole release status, work release status, and type of crime for which presently imprisoned. Rhodes (1989) studied the effects of the following risk variables separately on the probability and the timing of rearrest of inmates released from prison: race, gender, age at time of release from prison, number of prior convictions without imprisonment, heroin or opiate dependence, employment status, prior imprisonment, supervision status at the time of the sample imprisonment, and type of crime for which presently imprisoned. Visher and Linster (1990) studied the effects of the following risk variables on the time until pretrial rearrest for offenders released on their own recognizance: felony status of sample arrest, most serious initial arrest charge, number of positive drug test results, supervison status at the time of the sample arrest, employment status, education, age at arrest, number of prior convictions, and treatment intervention status.

therefore, be assessed, even in a preliminary way. As such, we are unable firmly to tap into these studies.

Schmidt and Witte (1988) and Rhodes (1989) have presented the most comprehensive reviews of the technical and practical aspects of multivariate failure time techniques used in criminological and cjs applications. Schmidt and Witte have apparantly developed the most general analytical strategy for using failure time techniques, examining the effects of risk variables separately on the probability and the timing of reimprisonment; their research specifically examined the effect of risk variables both on whether a released prisoner would eventually be reimprisoned and, given that this would occur, on the timing of the reimprisonment.[41] However, like all of the earlier failure time studies, which neglected risk variables, none of the more recent multivariate studies separately examined violent criminals and their violent crimes.

Visher's and Linster's study seems to be the only one to have employed the time of the instant arrest as the starting point for measuring the time until rearrest. Their study, which examined the antecedents of pretrial failure for offenders released on their own recognizance at arraignment, did not, however, concentrate on the violent offenders nor on these offenders' successive arrests for violent crimes.

While none of these earlier studies expressly dealt with arrests for criminal violence, they nevertheless supported several decisions made about the organization of the present study: (1) examining diverse parametric distributions, (2) incorporating risk variables, and (3) assessing prediction

---

[41] The authors hypothesized that those risk variables which influenced whether a person eventually returned to prison are different from those risk variables which influenced when that return would happen.

accuracy through a split-sample approach. First, because of their different foci and research strategies, these studies make it clear that one must examine diverse parametric failure time distributions each time one applies failure time techniques to different types of criminal behavior and/or cjs decision making. The possible applications of these techniques are much too diverse for the researcher simply to assume that a parametric distribution which performed admirably in one context will perform equally admirably elsewhere.. That stated, there is, however, some accumulating evidence that persons who have been apprehended by the cjs are at greatest risk of reinvolvment with the cjs fairly soon after they have exited from it (e.g., release to parole), and that this risk then begins to subside over time. Curvilinear failure time distributions, exhibiting either initially high or quickly increasing hazard rates which then quickly subside, would seem especially worthwhile exploring in the present context. Because this study is the first one to apply failure time methods to violent crime in the ways previously described, we do not want prematurely to rule out from consideration any plausible distributions, even those recommended by studies only peripherally related to the present one. For this reaon, we have elected to explore a wide assortment of distributions. Second, risk variables need to be incorporated into the parametric models in order to take into account individual differences in the probability and timing of rearrest. Parametric models which have included risk variables have tended more accurately to describe patterns in reinvolvement with the js than those parametric models which have not done so. For this reason, we used risk variables. Third, general prediction research has consistently hammered home the simple lesson that a statistical model which predicts future criminal behavior reasonably well in one group will not predict as well in another group (i.e., prediction shrinkage). Model validation, gauged in the present study by the level of predictive accuracy, was therefore critical to evaluate, and it was a primary concern of ours.

While there are no multivariate failure time studies which specifically identify risk variables increasing the probability of rearrest for serious violent crimes and/or decreasing the time until these rearrests, studies of general criminal careers provide some potentially useful leads in these regards. Considering only information which can be obtained from official records, the mainstay of the present study, offenders with lengthy, serious, and recent criminal records seem to be at greatest risk of continuing their criminal careers and of committing future crimes at high rates (Rhodes 1989). Several other risk variables have been found to be related to the future rate of individual criminal involvement; the type of first crime, the prior individual crime rate, and an early age at first criminal involvement.[42] If those risk variables which influence general criminal careers also influence the violent portion of these careers, then these variables may likewise aid in predicting the probability and timing of rearrest for violent criminal behavior. Whenever these risk variables were available, these and kindred ones were included in the present study.

---

[42] Farrington, D. P. 1987. Predicting individual crime rates. In Prediction and Classification: Criminal Justice Decision Making, special issue of Crime and Justice, A Review of Research, vol. 9, ed. D. M. Gottfredson, M. Tonry, 53-101. Chicago, IL: Univesity of Chicago Press, p. 94.

WHAT COMES NEXT?

The next chapter develops the prediction tools. First, unitary- and split-population parametric models without risk variables were estimated. These models were then compared at each point in the arrest sequence and across all points in the arrest sequence using the 1958 birth cohort construction sample and the total 1945 birth cohort sample. Second, the unitary-population parametric models with risk variables were estimated at each point in the arrest sequence using the 1958 birth cohort construction sample.[43] Third, based on those models estimated in the first two steps, the best parametric model was selected at each point in the arrest sequence. This comparative assessment enabled us to sort out the relative strengths of the different classes of parametric models (i.e., unitary- versus split-population) and to isolate which risk variables influenced the probability and timing of rearrest.

[43] This could not be done for the split-population models because of problems with algorithm convergence. See Maltz, M. D. 1984. Recidivism. Orlando, FL: Academic Press.

Chapter 3

THE RISK OF REARREST FOR VIOLENT CRIMES: WHAT TIMES ARE THE MOST RISKY? HOW STEEP ARE THE RISKS? WHAT VARIABLES INFLUENCE THE RISKS?

HOW LONG UNTIL REARREST? WHAT DO THE OBSERVED PERCENTILES SHOW?

The 1945 Birth Cohort

We began the statistical leg of the study by asking a simple question: How quickly, on the average, were the birth cohort subjects rearrested for their violent crimes? As a first step toward answering this question, we examined the overall (i.e., unconditional) observed (i.e., empirical) rearrest functions of the birth cohort subjects at successive arrest transitions.[1] The rearrest function yielded the percentage of subjects who were rearrested by successive points in time, for example, by the end of the first month, by the end of the second month, and so on, until exposure to rearrest terminated at the end of the juvenile or young adult ages.

This analysis highlighted some general aspects of the cohort subjects' rearrest risks and timing, which, in turn, established some useful departure points for later analyses. Examination of the overall rearrest function enabled us to identify with greater clarity the comparative impacts of selected risk variables on the overall trajectory of rearrest risks. Similarly, examination of the observed rearrest function, computed using common life-table methods, served as a baseline against which to compare the estimated (i.e., parametric) rearrest and hazard functions.

---

[1] In this study, the rearrest function is equivalent to the failure function discussed in Chapter 2.

Among the many risk variables which were examined, race was clearly the most troubling. This variable also happened to illustrate, dramatically at times, the great variability in rearrest risks and timing which were undergone by different birth cohort subjects. To stimulate sensitivity to this and other sources of variability in rearrest risks, we elected to incorporate this one variable into this initial analysis.

One way to grasp the main contours of the rearrest function is to flag the month by which a specific percentile (i.e., percentage) of the subjects had been rearrested. The cell entries in tables 3.1-2 list for the 1945 and 1958 birth cohort subjects, respectively, the months by which 10 percent, 25 percent, 50 percent, and 90 percent of the subjects were rearrested. These percentiles, obtained from the observed rearrest function, are presented for each arrest transition and race, for both juveniles (ages 10 to 17) and adults (ages 18 to 26).

First consider the 1945 birth cohort subjects. Among the total sample of juveniles at the first arrest transition, 10 percent were rearrested by the end of the 12th month, and 25 percent by the end of the 35th month. (The computation of the observed rearrest function indicated that 50 percent of the subjects had not been rearrested by the end of the juvenile period, by their 18th birthdays. Thus, the entry "not applicable" appears in the 50th-percentile column.) But, just one arrest transition later, subjects were rearrested at a much faster rate: 10 percent by the end of the first month and 25 percent by the end of the 16th month, representing sharp drops of one-twelfth and one-half, respectively. Were information available for more advanced arrest transitions, would this pattern in progressively more compressed rearrest times continue? The 1958 birth cohort data, to be discussed shortly, indicate that the answer to this question is probably yes.

Blacks dominated the total group of 1945 birth cohort subjects, during both the juvenile and adult periods and at each arrest transition during these time periods. To get a clearer picture of the comparative, and quite disproportionate, observed rearrest risks sustained by blacks and whites, rearrest functions were computed separately for each group. Black subjects were at much greater risk of rapid rearrest than white subjects. At the first juvenile arrest transition, ten percent of the blacks were rearrested within one year (11 months) in contrast to just slightly below six years (70 months) for whites; 25 percent of the blacks were rearrested in slightly more than two and one-half years (31 months), whereas whites bordered on six years (71 months) (table 3.1). The finding for whites of just a one-month difference between the 10th and 25th percentiles suggests that these subjects experienced predominantly (low) early rearrest risks. (This is reflected by the observed monthly hazard rates presented later in table 3.3.)

How do rearrest risks and timing in the 1945 birth cohort change as subjects advance from the juvenile to the young adult periods? Unfortunately, the reliability of the comparisons between these age periods was somewhat limited because there were relatively few adult subjects. Despite this limitation, the comparisons were helpful in pinpointing potentially important patterns in rearrest risks. We considered the 1945 birth cohort subjects' first arrest transitions in both the juvenile and young adult age periods, the only comparison that could be made. Adults were rearrested more rapidly than juveniles, regardless of whether they were black or white (table 3.1). The quickened pace of adult rearrests in comparison to juvenile rearrests ranged between about one-third and two-thirds at the 10th and 25th percentiles.

The 1958 Birth Cohort

The patterns in observed rearrest times exhibited by the 1958 birth cohort subjects mirrored those of the 1945 birth cohort subjects in two key respects: greater proportions of subjects were rearrested more rapidly as the arrest transition notched higher, and blacks were rearrested more rapidly than whites. Results failed to run parallel to one another, however, in one key respect: subjects in the 1958 birth cohort were rearrested less rapidly as adults than as juveniles.

Consider the juvenile patterns. Among the total group of 1958 birth cohort subjects, at the first arrest transition, 10 percent were rearrested within 4 months, 25 percent within 15 months, and 50 percent within 57 months (table 3.2). The length of time until each percentile was reached decreased almost uniformly with each successive arrest transition, and this decrease was clearest at the later, 50th percentile. The sharp differences across arrest transitions can easily be seen by comparing the two bracketing transitions: by the fifth transition, the 10th percentile was reached in less than one-eighth the time than at the first transition (less than one-half month versus 4 months), the 25th percentile in one-fifteenth the time (1 month versus 15 months), and the 50th percentile in about one-eleventh the time (5 months versus 57 months).

Black juveniles and adults were rearrested at much faster clips than were white juveniles and adults. Witness the divergence in each age period at the first arrest transition (the only one permitting a reliable comparison).

At the first juvenile transition, 10 percent of the black subjects were rearrested in one-fourth the time as white white subjects (3 months versus 12 months), 25 percent were rearrested in one-fifth the time (13 months versus 64 months), and 50 percent were rearrested in two-thirds the time (48 months versus 71 months). During young adulthood, this lopsided pattern was virtually replicated: 10 percent of the blacks were rearrested in one-fourth the time as whites (4 months versus 12 months) and 25 percent were rearrested in roughly one-fifth the time (20 months versus 88 months)

In contrast to the pattern displayed by the 1945 birth cohort, the 1958 birth cohort subjects were rearrested at a slower pace when they were adults than when they were juveniles, regardless of the arrest transition or the subject's race. To see this, look at the pattern for the total group. At the first transition, 10 percent of the adults were rearrested within 5 months in comparison to 4 months for the juveniles. This slim one-month difference in percentile times widened at the next milestone percentile: 25 percent of the subjects were rearrested within 27 months as adults in comparison to about one-half that amount (15 months) as juveniles. The gap between juveniles and adults at the first transition broadened by the fifth transition. It took twice as long for 10 percent of the adults to be rearrested at this final transition than juveniles (less than one-half month versus one month), four times as long for 25 percent to be rearrested (one month versus 4 months), and more than five times as long for 50 percent to be rearrested (5 months versus 26 months).

One can explain the reversal in the timing of rearrests across the juvenile and young adult ages in the two birth cohorts in two quite different ways: artifact and fact. On the one hand, the reversal might be an artifact of the relative unreliability of the thinner 1945 birth cohort data. Analyses confirmed the presence of this unreliability and how it weakened comparisons across age periods within the 1945 birth cohort and within the adult period across the two birth cohorts.[2] The reversal may, therefore, be more apparent than real. On the other hand, the reversal might be real, resulting, for example, from a more punitive response by the cjs in more recent years. Perhaps the 1958 birth cohort subjects were given more frequent and stiffer jail and prison sentences, lengthening the time until each rearrest by an amount equal to the time spent in confinement. (This is a time displacement effect.) It is impossible, however, to determine with certainty the correct alternative, and, in fact, the truth may lie somewhere between them. Despite this uncertainty, one thing is quite clear. Even if the 1958 birth cohort subjects were, indeed, at greater risk of being confined and of receiving and serving longer sentences than were the 1945 birth cohort subjects, the 1958 birth cohort subjects were still a very persistent lot when it came to rearrests; for example, at the fourth and fifth transitions, fully one-half of them were rearrested, within roughly two and two and one-half years.

Comparisons across Birth Cohorts

The most reliable cross-cohort comparisons, based on 30 or more subjects, could be made for: (1) both blacks and whites at the first juvenile

---

[2] The observed adult rearrest functions for the 1945 birth cohort subjects exhibited relatively high standard errors, producing wide confidence bands around the monthly, cumulative proportions of rearrested subjects. These confidence bands were often so wide that it was impossible to reject the possibilty that the contrasts in percentile times across the juvenile and adult periods of the 1945 birth cohort subjects were significantly different. The same uncertainty arose for comparisons of the juvenile periods in the two birth cohorts.

arrest transition, (2) blacks at the second juvenile arrest transition, and (3) blacks at the first adult arrest transition (tables 3.1-2). The juvenile comparisons were quite clear; the 1958 birth cohort subjects were rearrested more rapidly. For example, at the first arrest transition, 10 percent of the blacks in the 1958 birth cohort were rearrested within 3 months compared to 11 months for blacks in the 1945 birth cohort; in turn, 25 percent of the blacks in the 1958 birth cohort were rearrested within 13 months compared to 31 months for blacks in the 1945 birth cohort. A similar, but overall less pronounced, divergent pattern appeared for whites at the first juvenile arrest transition: 10 percent of the 1958 birth cohort subjects were rearrested in one-six the time it took the 1945 birth cohort subjects (12 months versus 70 months); but, once the 10th percentile time had been passed, the rearrest pace began to slacken for the 1958 birth cohort subjects, resulting in a just slightly shorter time until their 25th percentile was reached (64 months versus 71 months). The divergent patterns in rearrest timing during the juvenile period may presage later difficulties in making accurate individual- and aggregate-level predictions of rearrests across the two birth cohorts.

Adult comparisons, limited to the black subjects at the first arrest transition, indicated a virtual parity in rearrest timing: 10 percent of the black subjects in both birth cohorts were rearrested within 4 months and 25 percent were rearrested within 20 months. We will have the opportunity later to see whether this consistency presages reasonably accurate individual- and aggregate-level predictions across birth cohorts.

THE OBSERVED HAZARDS OF REARREST: SOME ILLUSTRATIVE PATTERNS WITHIN BIRTH COHORTS

One can think of each sequence of rearrest percentiles, obtained from the rearrest function, as having been generated by a corresponding hazard function. This is certainly reasonable to do because, after all, the rearrest and hazard functions are mathematically equivalent. One's research purposes determine which function to stress. If one is mainly interested in the timing of rearrests, the rearrest function is to be preferred. If one is mainly interested in the sequence of rearrest risks which produced the rearrest times, the hazard function is to be preferred. We now focus on the hazard function because it gets at the issue of rearrest risks.

Determining the shape of the observed, overall hazard function is an important priority in planning a productive study strategy and in developing rational and responsive public policies. On the one hand, if rearrest risks are generally highest during the period just after an arrest, the time until rearrest will be commensurately short. One will need to select parametric distributions which are able to reflect this pattern in high early risks; by implication, js interventions might usefully be marshalled and delivered during this early period. On the other hand, if the highest rearrest risks generally occur well after arrest, the time until rearrest will be commensurately long. One will need to select parametric distributions which are able to reflect this pattern in high later risks; by implication, js interventions might temporarily be held in reserve and then initiated at these later times. Tracing the ebb and flow in the magnitude of the hazard function is, therefore, both scientifically and practically important.

The 1945 and 1958 Birth Cohorts

a. Patterns in Juvenile and Adult Rearrest Risks

We moved to this next leg of the analysis by again asking a simple question. Precisely when, on the average, were the birth cohort subjects at greatest risk of being rearrested? To answer this question, we present in tables 3.3-3.5 the observed hazard rates for the 1945 and 1958 birth cohort subjects during their juvenile (ages 10 to 17) and young adult (ages 18 to 26) years. Hazard rates are presented for each month up to the 96th month in each age period. This provided coverage of the entire eight-year juvenile period and the first 96-month block of the 108-month long young-adult period. We decided for several reasons to calculate hazard rates for each month of a 96-month span in each age period: first, to ensure that the time intervals were brief enough to register noteworthy and abrupt changes in rearrest risks; second, to produce the longest possible duration of exposure to rearrest risks during each age period; and, third, to create comparable durations of exposure to rearrest risks in the two age periods.[3] (These are the reasons why, as one might have noticed, the adult age period examined in the earlier discussion of the arrest function was defined through age 26 rather than age 27, even though data were available through age 27.) The hazard rates in tables 3.3-5 were broken down by arrest transition and race.

Visual scanning of the observed hazard rates can sometimes result in uncertain and ambiguous interpretations of the patterns in rearrest risks, especially when comparing rates across different groups (e.g., juvenile versus

---

[3] The hazard rates were commonly so low during the final 12 months of the adult period that little information was lost by their exclusion.

adult). The computation of these rates, although not too complex, is nevertheless not simple. Fairly involved censoring patterns sometimes arise, injecting a level of complexity into the analysis which visual scanning cannot easily grasp. However, at this point in the study, we are interested only in pinpointing some broad features of the observed hazard rates, which will later serve as guideposts in assessing how well the estimated parametric distributions fit the observed data. For this reason, despite some limitations, we first visually scanned the observed hazard rates in order to uncover what we unambiguously could about them. As it turned out, the yield was not trivial.

Some clear and consistent patterns appeared across birth cohorts, arrest transitions, age periods, and races: first, rearrest risks were highest during the period immediately following arrest (the <u>dispersion</u> issue) and, second, the rearrest risk in any single month was generally quite low, even when that risk was at its zeneith (the <u>magnitude</u> issue). Some other risk patterns were also observed, but these were noted earlier in the discussion of the rearrest function: risks increased as the arrest transition advanced, reflected by more rapid rearrest times at later transitions; blacks were at greater risk than whites, also reflected by their more rapid rearrest times; and, based upon the more reliable juvenile comparisons, subjects in the 1958 birth cohort were at greater risk than subjects in the 1945 birth cohort, once again reflected by their more rapid rearrest times. To avoid rehashing earlier findings, we will only elaborate the first set of patterns relating to the dispersion and magnitudes of the hazard rates.

Overall, rearrest risks tended to be highest during the months immediately following arrest, roughly through the 12th to 18th months, regardless of birth cohort, age status, arrest transition, or race. For example, in the total group of 1945 birth cohort subjects, the highest hazard rates at the first two juvenile transitions occurred during the first month of each transition (.031 and .113), and the highest hazard rate at the first adult transition occurred during the second month (.042) (table 3.3). After the 18th month, the percentage of months exhibiting zero or negligible risks (probability < .005) increased substantially. Witness what happened at the first juvenile arrest transition for the total group of subjects: over the first 18 months, just one-third of the months registered zero or negligible rearrest risks (33 percent); over the next 18 months, the proportion more than doubled (78 percent); and, over the final 60 months, nearly every month had a zero or negligible risk (98 percent). A similar pattern appeared at the second juvenile arrest transition (first 18 months, 50 percent; second 18 months, 78 percent; final 32 months, 97 percent) and at the first adult transition (first 18 months, 44 percent; second 18 months, 66 percent; final 60 months, 88 percent).[4] Clearly, then, if a 1945 birth cohort subject were to be rearrested, he quickly confronted that risk. This pattern in the 1945 birth cohort was repeated regardless of whether we disaggregated the hazard rates by age period or race. This same pattern appeared in the 1958 birth cohort, also regardless of whether we disaggregated the hazard rates by age period or race.

Despite the high magnitudes of the hazard rates during the initial months in comparison to later months, these rates were modest in <u>absolute</u> terms, even when at their steepest--.031, .113, and .042 (table 3.3). We have

---

[4] At the second juvenile arrest trransition, the longest that subjects were exposed to rearrest was 68 months. Thus, the final exposure period was 32 months.

seen that these modest risks during the initial block of months, followed by even more modest risks during later months, were still able to exact sizable cumulative tolls over time in rearrested subjects, sometimes amounting to as much as 50 percent (tables 3.1-2). Whether these high base-rate rearrest transitions foretell greater predictive accuracy than the more commonly crossed (in the general crime prediction research) highly skewed, low base-rate rearrest transitions will be examined shortly. Remember the old saw that a low base rate usually condemns to failure high predictive accuracy because most risk variables tend to describe both those subjects who are rearrested and those who are not, thereby, failing to discriminate between the two groups. We will see whether, with these data, high base rates tell the lie to this saw.

It is comforting to see that the visual evidence of a decreasing hazard function is supported by statistical evidence. Three of the parametric distributions chosen for analysis in this study are characterized by hazard functions which can assume more than one trajectory, including a decreasing trajectory: the loglogistic, Weibull, and Gompertz. The hazard function decreases if the "shape" parameter of the Weibull is less than 1 (or greater than 1 if the extreme Weibull is used); if the shape parameter of the loglogistic is greater than 1; and if the shape parameter of the Gompertz is less than 0 (i.e, negative). In virtually every case, regardless of the birth cohort, the value of the shape parameter of each distribution indicated a decreasing hazard rate, and this value was usually statistically significant.

STATISTICAL MODELING OF THE OBSERVED HAZARD AND REARREST RATES: WHICH DISTRIBUTIONS LOOK BEST?

What Do the Loglikelihoods Show?

In view of the findings so far, we can begin to speculate about what form of parametric distribution might most closely match the observed distributions of rearrest times, as reflected by their corresponding hazard functions. Recall that the obsserved hazard functions generally displayed the highest rearrest risks just after the immediate arrest and then decreased thereafter. Parametric distributions whose hazard functions either rise sharply and then subside (e.g., loglogistic, lognormal) or simply start high and then subside (e.g., all split population distributions, mixed exponential, negative Weibull) are all plausible contenders. Based on these considerations, which one(s) looked best?

One way to assess the merits of rival parametric distributions is to compare their loglikelihood statistics. The less negative a distribution's loglikelihood, the better the overall match of that distribution to the observed data. While the use of loglikelihoods does not entail an exact statistical test of how well the distributions matched the observed data, because all of the distributions are not formally nested within a single parent distribution against which they can be sequentially compared, it is nevertheless a quite useful first step toward assessment. One informed rule-of-thumb suggests that the introduction of a parameter should "buy" a decrease in three loglikelihoods.[5] Otherwise the decrease is can be viewed as

[5] Maltz, M. 1991. "Survival Fitting and Analysis Software for Industrial, Biomedical, Correctional and Social Science Applications." Chicago, IL: University of Chicago at Chicago Circle.

artifactual and, therefore, irrelevant. Recall that the principle of economy, discussed earlier in relation to evaluating the merits of different parametric distributions, suggests that, all other things being equal, including their loglikelihood statistics, the distribution with the fewest parameters (i.e., the simplest one) is to be preferred. We adopted this principle.

Tables 3.6-8 present the loglikelihood statistics for the various parametric distributions. The tables also indicate, in the second column, the number of parameters characterizing each distribution. For the reader's review, and to prepare for later analyses, we have also entered in the column after each loglikelihood statistic in these tables the overall percentage of birth cohort subjects who were estimated by the indicated distribution to be eventually rearrested if given an unlimited amount of time exposed to the risk of rearrest. This estimated overall rearrest rate was useful when compared to the corresponding observed overall rearrest rate.

How well did the rival parametric distributions stack up against one another? There were two main findings in this regard: first, there was a single clear loser and, second, there were no clear winners. The exponential distribution, which asserts a constant risk of rearrest over time, was the clear loser. In virtually every comparison, the exponential distribution's loglikelihood statistic was the most highly negative. And, if it's loglikelihood statistic was not the most highly negative, it fell among those which were the most highly negative. This consistent finding of poor performance reflected the nonconstant, mostly declining, patterns in the observed hazard rates noted earlier; the exponential distribution was simply unable to represent the nonconstant risk patterns exhibited by the observed rearrest times.

The other, just as unmistakable, pattern in loglikelihood statistics was the virtually equivalent overall performance of nearly all of the other parametric distributions: no single distribution topped the field in both birth cohorts, age periods, and race groups. For example, partitioning the birth cohort subjects into a segment which would be rearrested and a segment which would not, as is done by the split-population distributions, hardly improved upon the corresponding unitary-population distribution's capacity to match the observed data patterns. Overall, across arrest transitions in the two birth cohorts, the difference in loglikelihood statistics between <u>any</u> two distributions, with the exception of the exponential distribution, generally clustered between one and three. Twenty-three of the 31 arrest transitions represented in tables 3.6-8 exhibited differences in loglikelihoods which fell into this range.

The most frustrating aspect of all this is that, with this information, we cannot now tell whether the distributions were equally good or equally bad in matching the observed data. This assessment will have to await the prediction applications. The matter will then be decided on the practical grounds of predictive accuracy.

<u>What Do the Estimated Rearrest Percentiles Show?</u>

We are not yet able to assess firmly the comparative merits of the rival unconditional parametric distributions, even after having jointly used as assessment criteria their loglikelihood statistics and economy of representation (i.e., simplicity as reflected in the fewest number of parameters). Most of the distributions appeared to perform about equally well. Whether this logjam can be pried apart by introducing risk variables, turning the unconditional analysis into a conditional analysis, remains to be seen. However, before we introduce the risk variables, we need to know some additional basic things about how well the different parametric distributions matched the observed distributions. To get a better handle on the extent of these matches, we looked at how well the parametric rearrest functions described the corresponding observed rearrest functions.

As a way to describe concisely the degree of match between the parametric and observed distributions, we returned to the percentile tables, but this time expanding them to display the month by which a specified rearrest percentile was <u>estimated</u> to be reached by a parametric distribution. Table 3.9 presents these estimates for the 1945 birth cohort subjects for each of the 10 distributions selected for examination; tables 3.10-3.15 present analogous estimates for the 1958 birth cohort subjects. The observed rearrest percentile times, discussed earlier, are listed again, in the first row, to make comparisons easier.

It was unusual to find that 90 percent of the subjects had been rearrested within the eight-year juvenile time span and the nine-year young adult time span, regardless of the birth cohort, arrest transition, and race. Comparisons between the observed and estimated percentiles were, therefore, restricted to the limited, lower range falling between 10 and 50.

For each <u>observed</u> rearrest percentile time which could be <u>calculated</u>, the parametric distributions (excepting the already discredited exponential distribution) generated corresponding estimated percentile times which clustered together, usually within about six-to-eight months of one another, and which did not usually differ by more than six-months from the observed percentile time. These differences are not very great. For example, look at table 3.9, at the observed 25th percentile for the first arrest transition. The observed time was 35 months (first row, second entry). The rearrest percentile times estimated by the parametric distributions, listed below the observed percentile time in the same column, did not differ from one another by more than three months, and none of these estimates differed from the observed rearrest percentile time by more than four months. This pattern appears elsewhere in both this and the other tables.

The one noteworthy exception to this close clustering pattern appeared for the white 1945 birth cohort subjects, when they were both juveniles and young adults, probably because of the more limited reliability of the data. Except for this anomaly, the overall comparability of the various parametric distributions retells, then, in a new way, the story previously told by the loglikelihood statistics: the parametric distributions were largely indistinguishable in their fitness in matching the observed data, at least as far as can be determined at this point in the analysis. And, reiterating a point made earlier in this regard, we do not yet know whether this comparability foretells uniformly accurate or inaccurate matching of the rearrest risks and timing.

The rearrest and related hazard functions of the selected parametric distributions diverge mainly in their right-hand tails. In the present context, the right-hand tail represents the later times at which the birth cohort subjects were exposed to rearrest risks. This divergence in the tails can be clearly seen with these data by reviewing the estimated 90th rearrest percentile times, which often sharply differed across distributions. The lognormal distribution, for example, has a very long thin tail, resulting in some of the largest estimates of the most advanced percentile times. Look at this divergence at the first juvenile arrest transition of the 1958 birth cohort subjects (table 3.10). The lognormal distribution estimated that 90 percent of the subjects would be rearrested within about 1,000 months (in more than 83 years!), in comparison to the next highest estimate of 773 months (nearly 65 years, also startlingly high!), produced by the loglogistic distribution. The lowest estimates were produced by the exponential and mixed exponential distributions (nearly 13 years--154 months--and nearly 18 years--213 months--respectively). The dispersion in estimated percentile times is always greatest at the highest percentile benchmarks.

If the birth cohort subjects had been observed for a much longer period of time, we might now be better able to assess the relative capacities of the parametric distributions to match the observed rearrest patterns, by comparing them at the more advanced times. However, the comparatively brief observation periods entailed by this study (96 months for juveniles and 108 months for young adults) provided little empirical grounds for making such assessments.[6]

A PRACTICAL CONSIDERATION: ARE THE TIMES UNTIL REARREST SUFFICIENTLY SHORT TO BE USEFUL IN AN APPLIED SETTING?

We now ask a practical question: Were sufficiently large numbers of the birth cohort subjects rearrested for serious violent crimes rapidly enough to warrant trying to identify them? Think of what it would mean if subjects who had been arrested while they were juveniles for a first serious violent crime took, on the average, four to five years to be rearrested. These subjects would have had to have been initially arrested at ages 13 and 14 in order to

[6] Because the bulk of rearrests occurred fairly quickly after the initial arrests, additional follow-up time might have yielded few fresh insights.

have been rearrested by their eighteenth birthdays. If they had not been arrested for the first time at these young ages, it would be unlikely that they would be rearrested while they were still juveniles, placing them beyond the jurisdiction, if not interest, of jjs officials.

The rearrest patterns displayed by the 1958 birth cohort answer the above question. Consider table 3.10, which presents the overall rearrest percentile times for the 1958 birth cohort subjects when they were passing through their juvenile years. By only the second arrest transition, each of the parametric distributions estimated a 50th percentile time in the neighborhood of just 18 months; and, by the fifth arrest transition, each of these estimates decreased to about six months. These data indicate, then, that an ample proportion of the subjects were rearrested with sufficient dispatch while they were juveniles to warrant their identification. A similar pattern appeared during the young adult years (table 3.13). Introducing risk variables into the analysis would almost certainly divide the birth cohort subjects described by tables 3.9-15 into subgroups which had very different overall rearrest times. The most noteworthy group from a public protection perspective are those who were rearrested quickly. We now look at which risk variables were related to a quickened pace of rearrest.

THE PREDICTION MODELS: RESULTS FROM THE MULTIVARIATE FAILURE TIME REGRESSIONS

Several noteworthy patterns emerged from the previous analysis of the overall observed data and their parametric representations: progressively shorter times between arrests at successively more advanced arrest transitions, higher risks of rearrest just after the initial arrest in an arrest transition, differences between juvenile and young adult rearrest risks (which shifted across the two birth cohorts), higher rearrest risks incurred by the black birth cohort subjects, and more pronounced rearrest risks during the juvenile period sustained by the more contemporary 1958 birth cohort. Can these patterns be explained in a consistent fashion--across arrest transitions, age intervals, and birth cohorts--by the risk variables selected for analysis by this study? To answer this question, multivariate failure time regression analyses were conducted, as described in the previous chapter.

For the reasons previously outlined, two sets of risk variables were created for use in the multivariaate analyses: (1) legally and ethically <u>permissible</u> and (2) legally and ethically <u>less permissible</u> and <u>impermissible</u>. These variables are listed in figure 2.1.

Figures 3.1-2 list the two risk variable sets according to the age interval (juvenile versus adult), arrest transition (1st through 5th), and race of the birth cohort subject (blacks and whites). (Race appears because only the black birth cohort subjects were arrested in large enough numbers to support the multivariate analyses at the higher arrest transitions.) An "X" indicates that the risk variable was included in the analysis of the designated racial group at the designated arrest transition. The broadest analyses used between 20 and 25 risk variables.

<u>Juvenile Arrest Transitions</u>

Table 3.16 summarizes the failure time regression results for the juvenile arrest transitions of the 1958 birth cohort construction sample. The table is split into five separate panels (A to E), each depicting a different arrest transition (1st through 5th). At each transition, five parametric failure time distributions were examined, which were listed as the column headings: proportional hazards, exponential, Weibull, loglogistic, and lognormal. (The proportional hazards model is actually semiparametric, although we will refer to it as parametric for ease of presentation.)[7]

Four risk-variable models were evaluated for each of the parametric distributions: (1) the baseline, or naive, model, which employed no risk variables (column "0"), (2) the legally permissible risk variable model (column "L"), (3) the legally-permissible-plus-race risk variable model (column "L + R"), and (4) the entire, or full, risk variable model (column "A"). This study design permitted us to see whether the introduction of additional risk variables having specific public policy significance (like race) discernibly improved the match of the model to the observed data. Because the models were successively broadened, each new model became a superset of the one immediately to its left. Formal tests of statistical significance could be conducted between the models because of the nested design.[8]

In order to compare the risk variable models estimated under each parametric distribution, we have reported in the first two rows of the table both the loglikelihood statistic and the number of risk variables corresponding to each model.[9] Whenever a broader model significantly improved the fit of a narrower model's match to the observed rearrest data (at p-val <

---

[7] The proportional hazards model was included because it provides robust estimates of the risk variables. The model served, then, as a good foil against which to compare results from the other models. The split-population models were excluded because they neither significantly nor consistently outperformed their corresponding unitary-population models. See table 3.16, note a, for additional discussion of the proportional hazards model.

[8] For further discussion of the model testing format, see table 3.16, note e.

[9] See table 3.16, note e, for more details about the testing procedure.

.05), that improvement was represented by an asterisk to the right of the loglikelihood statistic. This format helped us systematically to determine whether legally permissible risk variables were associated with the more compressed rearrest times, whether race was associated with the more compressed rearrest times; and, finally, whether any of the other legally less permissible or impermissible risk variables were related to the more compressed rearrest times.

The shape and scale parameters of the parametric distributions are also presented. These appear in the rows just below the loglikelihood statistic.[10] The shape parameter is especially important because it provides information about the curvature of the distribution's hazard function. When a shape parameter is followed by an asterisk, this indicates that the distribution's hazard function decreased with the passage of time.[11]

Because both the legally permissible risk variables and race are so central to js decision making and to public policy making in this area, these variables always appear in the table panels, regardless of whether they were statistically significant or not. This format drives home the message about their relative utility in js decision making. All other risk variables appear in the tables only when they were statistically significant.

We now turn to the analysis of the 1958 birth cohort subjects during their juvenile years, focusing on the total group at their first arrest transition (table 3.16, Panel A.1) Several noteworthy, broad patterns appeared and, moreover, these patterns generally reappeared at other arrest

[10] Dashes appear in these rows for the proportional hazards model because it is not characterized by these parameters.

[11] See table 3.16, note f, for more details about the shape parameters of the selected distributions.

transitions, during young adulthood, and in the 1945 birth cohort. First, the legally permissible risk variables failed to improve the match of the models to the observed rearrest times. (No asterisk appeared in the "-2 loglikelihood" row under column "L" for any parametric distribution.) Second, race was significantly related to the timing or rearrests; black birth cohort subjects were rearrested more quickly, expressed by the positive coefficient in column "L + R" under the proportional hazards formulation and the negative coefficient in column "L + R" under each of the other parametric distributions.[12] Third, the race effect remained intact even when it was challenged by all of the other risk variables. The race variable was uniformly significant in each "A" risk variable model. Fourth, few risk variables overall achieved statistical significance. Fifth, when a risk variable was significant in more than one risk variable model, the signs were the same and the magnitudes comparable. Sixth, the shape parameters of the Weibull and loglogistic distributions were consistently significantly greater than one, indicating that their hazard functions decreased over time. Seventh, the exponential distribution least accurataely matched the observed rearrest times, based on a comparison of loglikelihoods.

These patterns were largely repeated among the black and white birth cohort subjects at the first arrest transition, and among the black subjects at the second, third, and fourth arrest transitions (table 3.16, Panels A.2-3, B-D). (The fifth arrest transition, presented in Panel E, was excluded from the comparison because there were too few cases to examine the all-risk-

[12] See table 3.16, note a, for further discussion of how to interpret the coefficients of the parametric distributions.

variables model.) Also, scanning across arrest transitions, no risk variable was consistently related to quick rearrest times.

In order to check the generality of the above results, the analyses were replicated using the 1945 birth cohort subjects. Because there were fewer subjects in the 1945 birth cohort, cross-cohort comparisons were limited to the first two juvenile arrest transitions. Table 3.17 presents the results. The main patterns appearing in the 1958 birth cohort reappeared in the 1945 birth cohort: legally permissible risk variables were unrelated to the timing of rearrest, blacks were rearrested more quickly than whites, few risk variables consistently achieved significance across both parametric distributions and arrest transitions, and the shape parameters of the Weibull and loglogistic distributions indicated the presence of decreasing hazard functions.

Adult Arrest Transitions

Analyses identical to those discussed above were conducted for the adult arrest transitions of the 1958 birth cohort subjects. These analyses assumed that the birth cohort subjects were "reborn" as adult violent criminals--that their juvenile records were sealed and, therefore, not employed in official decision making. As a consequence, we only used information about each subject's prior adult criminal record.

Table 3.18 presents these results. Because adults were arrested in greater numbers than were juveniles, we were able to analyze blacks and whites separately at the first two arrest transitions. Several of the overall patterns observed during the juvenile years also appeared in the adult years. First, with the exception of the first and second arrest transitions which combined the black and white birth cohort subjects (Panels A.1 and B.1), few legally permissible risk variables were related to the quick rearrest times. Second, race was associated with rearrest timing at the first arrest transition in both the "L + R" and the "A" models, as it was at the first juvenile arrest transition, but it failed to maintain this association at the second arrest transition. Third, overall few risk variables achieved statistical significance, and those which did achieve significance did not do so consistently across arrest transitions. Fourth, the shape parameters of the Weibull and loglogistic distributions were usually significantly greater than one, indicating that the adult arrest transitions were also characterized by decreasing hazard functions. Fifth, the exponential distribution consistently matched the observed rearrest times worse than the other distributions. Sixth, none of the other distributions consistently performed better than the others in matching the observed rearrest times.

Augmenting the Analyses of the Juvenile and Adult Arrest Transitions

It was important to try to augment the multivariate analyses by increasing the reliabilities of the rearrest time and risk variables. Doing so would also give some idea of how sensitive the previous results were to alternative data specifications. Two options were directly available for increasing reliabilities. The first option, applying to the juvenile period, involved extending the rearrest exposure time from age 18 to 27. The period over which juvenile rearrest transitions were followed was simply extended by nine years. In all other respects (i.e., variables, scaling, design), the analysis remained the same. The second option involved lowering the age floor of the adult rearrest analysis. Those risk variables describing the birth cohort subjects' prior records were permitted to extend backward into the juvenile period. Thus, for example, a birth cohort subject's age at the first prior offense was no longer restrictively defined as the age at the first prior adult offense but rather as the age at the first prior offense overall, including juvenile offenses. This procedure enabled us to explore the merits of using juvenile records in adult criminal cases.

We first turned to the juvenile period, extending the time window through age 27. Tables 3.19-23 present some of the descriptive background for this analysis. Table 3.19 presents the observed rearrest time percentiles based on the nine-year time extension. The findings in table 3.19 can be compared to the findings in table 3.2, which presented analogous percentile time information with respect to the more restricted time window, through age 18. The comparison is straightforward and unsurprising. The upper percentile times increased due to the longer exposure to rearrest risks, which, quite simply, permitted more subjects to be rearrested at the older ages. The higher observed rearrest time percentiles are reflected in the higher rearrest time percentiles estimated by the parametric distributions (tables 3.20-21). To see this amplifying effect of increased exposure to rearrest risks on the estimated percentile times, one can compare table 3.20 to table 3.10 (blacks and whites together) and table 3.21 to tables 3.11-12 (blacks and whites separately). For the reader's review, we also present the observed hazard rates based on the extended observation time (table 3.22). These hazard rates display the now familiar pattern of high early rearrest risks which then decrease with the passage of time.

Does the additional exposure time aid in distinguishing the better from the worse parametric distributions? Yes, but in a limited way, at just some of the arrest transitions. The greater capacity to match the observed pattern in rearrest times provided by the additional exposure time can be seen by comparing the loglikelihood statistics of the parametric models for those models estimated using the shorter observation window (table 3.7) and those estimated using the longer observation window (table 3.23). What had been only slight differences in loglikelihoods before (table 3.7) showed up much stronger now (table 3.23), but mainly at the earlier arrest transitions. The main finding indicated the superiority of the split-population and mixed-population models in comparison to the unitary-population models.

Table 3.24 summarizes the multivariate regression results based on the extended observation time. The overall pattern in results remained intact. Although some additional risk variables were significant in the augmented analysis, mainly at the early arrest transitions, no clearcut new pattern emerged. With the exception of the first arrest transition (Panel A.1), legally permissible variables did not seem to be highly related to rearrest times. Race, on the other hand, was significantly related to rearrest timing (Panel A.1), as it had been in the earlier analsysis. The loglogistic distribution appeared to fare better than the others in matching the observed data. Also, the distributional shape parameters suggested that the failure times followed a decreasing hazard function.

Table 3.25 shows the multivariate results when the prior criminal record variables were created using both juvenile and adult information. Table 3.25 can be compared to table 3.18, which restricts the prior criminal records to reflect only adult activity. Note that the only tabular entries which changed across tables were those in column "A". Several patterns appeared. First, a few more variables were statistically significant in this augmented version of the data, mainly at the first two arrest transitions. Second, race again was significant at the first arrest transition, but it failed to remain significant at the second arrest transition. Third, no risk variables consistently appeared to be significant across arrest transitions. Fourth, the Weibull and loglogistic shape parameters indicated declining hazard rates, as they had done in the comparison analyses. Fifth, the loglikelihoods of these augmented parametric models (table 3.25) were lower than those of the corresponding nonaugmented models (table 3.18), indicating the greater explanatory utility of these variables.

Chapter 4

SUMMARY AND NEXT STEPS

So far, this study has developed sequential-prediction models to be applied to arrests for serious violent crimes. To develop these models, we used failure time regression techniques at each rung in the arrest chain of the construction sample of the 1958 Philadelphia birth cohort subjects. These failure time regression analyses involved examining five parametric distributions: proportional hazards, exponential, Weibull, loglogistic, lognormal. For each parametric distribution, four nested risk-variable models were investigated--a naive model of no risk variables, the legally permissible risk variables, the legally permissible risk variables plus race, and all risk variables. This analysis design enabled us to determine, for each parametric distribution, whether the legally permissible risk variables were related to a high risk of rapid rearrest, whether the race variable nullified the effects of significant legally permissible risk variables, and whether other risk variables were related to a high risk of rapid rearrest.

The failure time regression models indicated that legally permissible risk variables were not often associated with rearrest risks and timing, that race had a consistent effect, but at the first arrest transition, and that few risk variables overall were related to rearrest. We were unable to identify risk variables which were consistently significant across arrest transitions, age groups, and birth cohorts.

We are now conducting individual- and aggregate-prediction analyses based on these failure time regression findings. The best fitting sequential-prediction models estimated at each rearrest rung are being used to produce individual and aggregate predictions. These predictions are being applied to the validation sample of the 1958 birth cohort and to subsets of the 1945 birth cohort. The achieved levels of predictive accuracy will dictate the practical utility of the prediction models.

More work will need to be done to assess and enhance the prediction results, whatever their observed level of accuracy. This work might include: (1) expanding the types of criminal behaviors to be predicted as a way to increase the reliability of the outcome measure (e.g., property index crimes in addition to the violent index crimes), (2) expanding the pool of risk variables by using nonofficial data for the 1958 birth cohort subjects, (3) exploring new types of failure time models which specifically address repeated events like rearrests, and (4) examining failure time models which permit one to predict the type of rearrest (e.g., competing hazards). These options are being investigated.

Table 2.1

Birth Cohort Analyses by Age Interval

Birth Cohort	Number of Subjects	Age Interval		
		10-17	18-26	10-26
1945				
Total	9,945	X	-	-
Follow-up Sample	978	X	X	X
1958	13,160	X	X	X

Table 2.2

The Number of Birth Cohort Subjects Arrested and the Number of Times They Were Arrested for Violent Crimes by Birth Cohort and Age Interval

	Subjects		Arrests	
Birth Cohort	10-17	18-26	10-17	18-26
1945				
Total	360	--	435 (2)[a]	--
Follow-up Sample	25	74	31 (2)	104 (2)
1958				
Construction	759	911	1,244 (5)	1,516 (5)
Validation	324	393	563 (5)	639 (5)

a. The figure in parentheses is the number of arrest transitions used to calculate the number of times the birth cohort subjects were arrested for violent crimes. The number of arrest transitions listed here for a particular birth cohort and age interval equals the number of arrest transitions examined throughout this study for that birth cohort and age interval.

Table 2.3

The Number of Subjects Arrested by Birth Cohort, Arrest Transition, and Age Interval

Panel A: 1945 Birth Cohort

Arrest Transition	Total	Follow-up Sample	
	10-17	10-17	18-26
1	360	25	74
2	75	6	30

Panel B: 1958 Birth Cohort

Arrest Transition	Construction		Validation	
	10-17	18-26	10-17	18-26
1	759	911	324	393
2	262	325	128	140
3	124	157	62	67
4	62	83	28	35
5	37	40	21	18

Table 3.1

1945 Birth Cohort: Arrests for Violent Crimes--
Selected Observed Rearrest-Time (in Months)
Percentiles by Age Status, Race,
and Arrest Transition

Age Status and Race	1st (N)[a]	10	25	50	90	2nd (N)	10	25	50	90
Juveniles (Total Sample)										
Total	(360)	12	35	NA[b]	NA	(75)	1	16	NA	NA
Blacks	(302)	11	31	NA	NA	(72)	2	18	NA	NA
Whites	(58)	70	71	NA	NA	(3)[c]	--[c]	--	--	--
Adults (Follow-up Sample)										
Total	(74)	8	23	NA	NA	(30)	--	--	--	--
Blacks	(56)	4	20	55	NA	(26)	--	--	--	--
Whites	(18)	22	26	NA	NA	(4)	--	--	--	--

a. The number of birth cohort subjects at risk of rearrest.

b. The cell entry is not applicable because the percentile was not reached.

c. There were too few cases ($N \leq 30$) to compute the rearrest time percentile. White adults at the first arrest transition were exempted from this threshold in order to provide some comparative findings.

Table 3.2

1958 Birth Cohort: Arrests for Violent Crimes--
Selected Observed Rearrest-Time (in Months)
Percentiles by Age Status, Race, and
Rearrest Transition
(Construction Sample)

Age Status and Race	1st					2nd					3rd					4th					5th				
	(N)[a]	10	25	50	90	(N)	10	25	50	90	(N)	10	25	50	90	(N)	10	25	50	90	(N)	10	25	50	90
Juveniles																									
Total	(759)	4	15	57	NA[b]	(262)	2	7	18	NA	(124)	1	3	17	NA	(62)	1	4	10	45	(37)	<.5	1	5	NA
Blacks	(644)	3	13	48	NA	(245)	2	7	18	NA	(117)	1	3	17	NA	(59)	1	4	10	45	(36)	<.5	1	6	NA
Whites	(115)	12	64	71	NA	(17)[c]	--	--	--	--	(--)	--	--	--	--	(--)	--	--	--	--	(--)	--	--	--	--
Adults																									
Total	(911)	5	27	NA	NA	(325)	2	12	48	NA	(157)	4	12	39	NA	(83)	1	11	30	NA	(40)	1	4	26	NA
Blacks	(693)	4	20	NA	NA	(277)	3	13	47	NA	(137)	4	12	43	NA	(69)	1	10	30	NA	(34)	1	4	26	NA
Whites	(218)	12	88	NA	NA	(48)	1	6	NA	NA	(20)	--	--	--	--	(--)	--	--	--	--	(--)	--	--	--	--

a. The number of birth cohort subjects at risk of rearrest.

b. The cell entry is not applicable because the percentile was not reached.

c. There were too few cases (N $\leq$ 30) to compute the rearrest time percentile.

Table 3.3

1945 Birth Cohort: Arrests for Violent Crimes--
Observed Monthly Hazard Rates by Age Status,
Race, and Arrest Transition

	Juveniles (Total Sample)					Adults (Follow-up Sample)		
	Total		Blacks		Whites	Total	Blacks	Whites
Month	1st (N = 360)[a]	2nd (N = 75)	1st (N = 302)	2nd (N = 72)	1st (N = 58)	1st (N = 74)	1st (N = 56)	1st (N = 18)
1	.031	.113	.034	.102	.017	.014	.018	.000
2	.000	.015	.000	.000	.000	.042	.056	.000
3	.009	.015	.010	.016	.000	.014	.019	.000
4	.003	.000	.003	.000	.000	.015	.020	.000
5	.009	.000	.010	.000	.000	.000	.000	.000
6	.003	.016	.004	.016	.000	.015	.020	.000
7	.009	.000	.011	.000	.000	.000	.000	.000
8	.009	.000	.011	.000	.000	.000	.000	.000
9	.000	.032	.000	.033	.000	.015	.021	.000
10	.006	.000	.007	.000	.000	.031	.021	.057
11	.006	.000	.007	.000	.000	.000	.000	.000
12	.012	.000	.015	.000	.000	.016	.022	.000
13	.019	.017	.019	.017	.018	.016	.022	.000
14	.010	.000	.012	.000	.000	.000	.000	.000
15	.003	.017	.004	.017	.000	.000	.000	.000
16	.000	.018	.000	.018	.000	.000	.000	.000
17	.006	.000	.008	.000	.000	.000	.000	.000
18	.007	.018	.008	.018	.000	.033	.046	.000
19	.007	.019	.008	.019	.000	.000	.000	.000
20	.013	.000	.016	.000	.000	.000	.000	.000
21	.003	.020	.004	.020	.000	.034	.048	.000
22	.003	.000	.004	.000	.000	.018	.000	.061
23	.000	.000	.000	.000	.000	.018	.025	.000
24	.003	.020	.004	.021	.000	.000	.000	.000
25	.000	.000	.000	.000	.000	.037	.026	.066
26	.003	.021	.004	.022	.000	.019	.000	.073
27	.003	.000	.004	.000	.000	.000	.000	.000
28	.000	.000	.000	.000	.000	.000	.000	.000
29	.004	.000	.004	.000	.000	.000	.000	.000
30	.004	.000	.004	.000	.000	.000	.000	.000
31	.004	.000	.004	.000	.000	.000	.000	.000
32	.004	.000	.004	.000	.000	.000	.000	.000
33	.007	.000	.009	.000	.000	.000	.000	.000
34	.000	.000	.000	.000	.000	.000	.000	.000
35	.004	.000	.005	.000	.000	.040	.054	.000
36	.007	.000	.009	.000	.000	.000	.000	.000
37	.000	.023	.000	.023	.000	.021	.028	.000
38	.004	.000	.005	.000	.000	.000	.000	.000
39	.000	.000	.000	.000	.000	.000	.000	.000
40	.004	.000	.005	.000	.000	.021	.030	.000
41	.000	.000	.000	.000	.000	.000	.000	.000
42	.000	.000	.000	.000	.000	.022	.031	.000
43	.000	.000	.000	.000	.000	.000	.000	.000
44	.000	.000	.000	.000	.000	.000	.000	.000
45	.000	.000	.000	.000	.000	.023	.032	.000
46	.000	.000	.000	.000	.000	.000	.000	.000
47	.004	.000	.005	.000	.000	.000	.000	.000
48	.004	.000	.005	.000	.000	.000	.000	.000
49	.000	.000	.000	.000	.000	.000	.000	.000
50	.000	.000	.000	.000	.000	.024	.034	.000
51	.000	.000	.000	.000	.000	.000	.000	.000
52	.000	.000	.000	.000	.000	.000	.000	.000
53	.000	.000	.000	.000	.000	.000	.000	.000
54	.000	.000	.000	.000	.000	.000	.000	.000
55	.000	.000	.000	.000	.000	.024	.035	.000
56	.000	.000	.000	.000	.000	.000	.000	.000
57	.000	.000	.000	.000	.000	.000	.000	.000
58	.000	.000	.000	.000	.000	.000	.000	.000
59	.000	.000	.000	.000	.000	.000	.000	.000
60	.000	.000	.000	.000	.000	.000	.000	.000
61	.000	.000	.000	.000	.000	.000	.000	.000
62	.008	.000	.010	.000	.000	.000	.000	.000
63	.000	.000	.000	.000	.000	.000	.000	.000

Table 3.3 (cont.)

	Juveniles (Total Sample)					Adults (Follow-up Sample)		
	Total		Blacks		Whites	Total	Blacks	Whites
Month	1st (N = 360)	2nd (N = 75)	1st (N = 302)	2nd (N = 72)	1st (N = 58)	1st (N = 74)	1st (N = 56)	1st (N = 18)
64	.000	.000	.000	.000	.000	.000	.000	.000
65	.000	.000	.000	.000	.000	.000	.000	.000
66	.000	.000	.000	.000	.000	.000	.000	.000
67	.000	.000	.000	.000	.000	.000	.000	.000
68	.000	.000	.000	.000	.000	.000	.000	.000
69	.000	NA[b]	.000	NA	.000	.025	.037	.000
70	.000	NA	.000	NA	.000	.000	.000	.000
71	.004	NA	.000	NA	.021	.000	.000	.000
72	.000	NA	.000	NA	.000	.000	.000	.000
73	.000	NA	.000	NA	.000	.000	.000	.000
74	.000	NA	.000	NA	.000	.000	.000	.000
75	.000	NA	.000	NA	.000	.000	.000	.000
76	.000	NA	.000	NA	.000	.000	.000	.000
77	.000	NA	.000	NA	.000	.000	.000	.000
78	.000	NA	.000	NA	.000	.000	.000	.000
79	.000	NA	.000	NA	.000	.000	.000	.000
80	.000	NA	.000	NA	.000	.000	.000	.000
81	.000	NA	.000	NA	.000	.000	.000	.000
82	.000	NA	.000	NA	.000	.000	.000	.000
83	.000	NA	.000	NA	.000	.000	.000	.000
84	.000	NA	.000	NA	.000	.000	.000	.000
85	.000	NA	.000	NA	.000	.000	.000	.000
86	.000	NA	.000	NA	.000	.000	.000	.000
87	.000	NA	.000	NA	.000	.000	.000	.000
88	.000	NA	.000	NA	.000	.000	.000	.000
89	.000	NA	.000	NA	.000	.000	.000	.000
90	.000	NA	.000	NA	.000	.000	.000	.000
91	.000	NA	.000	NA	.000	.000	.000	.000
92	.000	NA	.000	NA	.000	.000	.000	.000
93	.005	NA	.006	NA	.000	.000	.000	.000
94	.000	NA	.000	NA	.000	.000	.000	.000
95	.000	NA	.000	NA	.000	.000	.000	.000
96	.000	NA	.000	NA	.000	.000	.000	.000

a. The number of birth cohort subjects at risk of rearrest.

b. The cell entry is not applicable because no birth cohort subjects were exposed to the risk of rearrest during this month.

Table 3.4

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Observed Monthly Hazard Rates by Race and
Arrest Transition
(Construction Sample)

	Total					Blacks					Whites
Month	1st (N = 759)[a]	2nd (N = 262)	3rd (N = 124)	4th (N = 62)	5th (N = 37)	1st (N = 644)	2nd (N = 245)	3rd (N = 117)	4th (N = 59)	5th (N = 36)	1st (N = 115)
1	.046	.071	.157	.157	.277	.051	.068	.157	.165	.286	.018
2	.017	.046	.069	.019	.037	.020	.049	.073	.020	.038	.000
3	.018	.040	.031	.020	.120	.022	.038	.033	.021	.081	.000
4	.031	.014	.076	.083	.139	.036	.014	.070	.088	.138	.000
5	.016	.047	.047	.022	.052	.020	.045	.050	.023	.052	.000
6	.015	.019	.000	.144	.000	.011	.021	.000	.154	.000	.036
7	.009	.030	.012	.027	.057	.007	.032	.013	.029	.057	.019
8	.014	.047	.050	.028	.063	.015	.051	.040	.030	.063	.010
9	.022	.062	.026	.058	.070	.027	.066	.028	.031	.070	.000
10	.008	.024	.000	.031	.080	.010	.026	.000	.032	.080	.000
11	.016	.018	.042	.066	.000	.020	.013	.044	.068	.000	.000
12	.007	.006	.014	.035	.000	.008	.007	.015	.037	.000	.000
13	.017	.019	.015	.000	.095	.016	.020	.016	.000	.095	.019
14	.017	.039	.015	.000	.121	.019	.043	.016	.000	.120	.010
15	.012	.007	.016	.038	.000	.015	.007	.016	.039	.000	.000
16	.002	.028	.000	.000	.000	.002	.022	.000	.000	.000	.000
17	.009	.022	.000	.040	.000	.011	.024	.000	.042	.000	.000
18	.009	.008	.033	.044	.000	.011	.008	.035	.045	.000	.000
19	.007	.008	.035	.000	.000	.009	.008	.037	.000	.000	.000
20	.009	.008	.000	.000	.000	.011	.008	.000	.000	.000	.000
21	.008	.016	.019	.048	.000	.009	.009	.020	.050	.000	.000
22	.008	.008	.000	.000	.000	.009	.009	.000	.000	.000	.000
23	.004	.017	.019	.000	.000	.005	.018	.021	.000	.000	.000
24	.002	.009	.020	.000	.000	.000	.009	.022	.000	.000	.010
25	.006	.018	.000	.056	.000	.007	.020	.000	.058	.000	.000
26	.008	.000	.000	.000	.000	.010	.000	.000	.000	.000	.000
27	.002	.010	.022	.000	.000	.002	.010	.023	.000	.000	.000
28	.002	.010	.000	.000	.000	.002	.011	.000	.000	.000	.000
29	.002	.000	.000	.000	.000	.002	.000	.000	.000	.000	.000
30	.010	.021	.000	.000	.000	.010	.023	.000	.000	.000	.010
31	.000	.012	.000	.000	NA[b]	.000	.013	.000	.000	NA	.000
32	.008	.012	.000	.074	NA	.010	.014	.000	.076	NA	.000
33	.008	.000	.000	.000	NA	.008	.000	.000	.000	NA	.011
34	.009	.000	.000	.000	NA	.011	.000	.000	.000	NA	.000
35	.002	.000	.000	.000	NA	.003	.000	.000	.000	NA	.000
36	.007	.000	.000	.000	NA	.008	.000	.000	.000	NA	.000
37	.004	.000	.000	.000	NA	.006	.000	.000	.000	NA	.000
38	.000	.000	.000	.000	NA	.000	.000	.000	.000	NA	.000
39	.005	.000	.000	.000	NA	.006	.000	.000	.000	NA	.000
40	.005	.000	.000	.000	NA	.006	.000	.000	.000	NA	.000
41	.000	.000	.000	.000	NA	.000	.000	.000	.000	NA	.000

Table 3.4 (cont.)

Month	Total 1st (N = 759)	Total 2nd (N = 262)	Total 3rd (N = 124)	Total 4th (N = 62)	Total 5th (N = 37)	Blacks 1st (N = 644)	Blacks 2nd (N = 245)	Blacks 3rd (N = 117)	Blacks 4th (N = 59)	Blacks 5th (N = 36)	Whites 1st (N = 115)
42	.002	.012	.000	.000	NA	.003	.016	.000	.000	NA	.000
43	.000	.000	.000	.000	NA	.000	.000	.000	.000	NA	.000
44	.000	.000	.000	.000	NA	.000	.000	.000	.000	NA	.000
45	.002	.000	.000	.000	NA	.003	.000	.000	.000	NA	.000
46	.000	.000	.000	.185	NA	.000	.000	.000	.191	NA	.000
47	.002	.000	.000	NA	NA	.003	.000	.000	NA	NA	.000
48	.002	.000	.000	NA	NA	.003	.000	.000	NA	NA	.000
49	.002	.000	.000	NA	NA	.003	.000	.000	NA	NA	.000
50	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000
51	.002	.000	.000	NA	NA	.003	.000	.000	NA	NA	.000
52	.002	.000	.000	NA	NA	.003	.000	.000	NA	NA	.000
53	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000
54	.003	.000	.000	NA	NA	.003	.000	.000	NA	NA	.000
55	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000
56	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000
57	.005	.000	.000	NA	NA	.007	.000	.000	NA	NA	.000
58	.003	.000	.000	NA	NA	.003	.000	.000	NA	NA	.000
59	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
60	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
61	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
62	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
63	.000	NA	.030	NA	NA	.000	NA	.032	NA	NA	.000
64	.006	NA	.000	NA	NA	.004	NA	.000	NA	NA	.012
65	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
66	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
67	.003	NA	.000	NA	NA	.004	NA	.000	NA	NA	.000
68	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
69	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
70	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
71	.003	NA	.000	NA	NA	.000	NA	.000	NA	NA	.017
72	.000	NA	.000	NA	NA	.004	NA	.000	NA	NA	.000
73	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
74	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
75	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
76	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
77	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
78	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	.000
79	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA
80	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA
81	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA
82	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA
83	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA
84	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA
85	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA
86	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA

Table 3.4 (cont.)

Month	Total 1st (N = 759)	Total 2nd (N = 262)	Total 3rd (N = 124)	Total 4th (N = 62)	Total 5th (N = 37)	Blacks 1st (N = 644)	Blacks 2nd (N = 245)	Blacks 3rd (N = 117)	Blacks 4th (N = 59)	Blacks 5th (N = 36)	Whites 1st (N = 115)
87	.000	NA	.000	NA	NA	.000	NA	.000	NA	NA	NA
88	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA
89	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA
90	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA
91	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA
92	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA
93	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA
94	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA
95	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA
96	.000	NA	NA	NA	NA	.000	NA	NA	NA	NA	NA

a. The number of birth cohort subjects at risk of rearrest.

b. The cell entry is not applicable because no birth cohort subjects were exposed to the risk of rearrest during this month.

Table 3.5

1958 Birth Cohort: Adult Arrests for Violent Crimes--
Observed Monthly Hazard Rates by Race and
Arrest Transition
(Construction Sample)

	Total					Blacks					Whites	
Month	1st (N = 911)[a]	2nd (N = 325)	3rd (N = 157)	4th (N = 83)	5th (N = 40)	1st (N = 693)	2nd (N = 277)	3rd (N = 137)	4th (N = 69)	5th (N = 34)	1st (N = 218)	2nd (N = 48)
1	.040	.054	.046	.101	.105	.047	.041	.045	.091	.125	.019	.133
2	.020	.047	.020	.083	.087	.018	.046	.023	.083	.069	.024	.049
3	.019	.028	.021	.000	.031	.020	.024	.016	.000	.036	.014	.051
4	.016	.014	.021	.000	.065	.019	.016	.024	.000	.077	.005	.000
5	.021	.029	.029	.015	.000	.028	.029	.033	.017	.000	.000	.027
6	.011	.030	.015	.045	.069	.013	.026	.017	.054	.041	.005	.056
7	.005	.015	.053	.000	.000	.003	.017	.062	.000	.000	.010	.000
8	.016	.015	.000	.016	.074	.020	.013	.000	.019	.087	.005	.029
9	.009	.008	.016	.016	.040	.007	.009	.009	.019	.000	.015	.000
10	.014	.012	.024	.000	.000	.019	.014	.019	.000	.000	.000	.000
11	.009	.020	.025	.016	.041	.012	.023	.019	.020	.047	.000	.000
12	.009	.008	.034	.033	.000	.009	.009	.019	.040	.000	.010	.000
13	.012	.016	.009	.052	.043	.016	.019	.010	.064	.050	.000	.000
14	.005	.008	.026	.036	.000	.007	.010	.030	.045	.000	.000	.000
15	.004	.013	.027	.019	.000	.002	.010	.031	.024	.000	.010	.000
16	.004	.004	.009	.000	.046	.006	.005	.011	.000	.053	.000	.000
17	.005	.009	.000	.019	.000	.006	.010	.000	.024	.000	.005	.000
18	.007	.013	.028	.000	.000	.009	.015	.011	.000	.000	.000	.000
19	.013	.004	.000	.040	.000	.015	.005	.000	.025	.000	.005	.000
20	.003	.018	.029	.000	.000	.004	.015	.033	.000	.000	.000	.031
21	.004	.014	.010	.021	.000	.002	.016	.011	.026	.000	.011	.000
22	.011	.009	.020	.021	.000	.012	.005	.023	.027	.000	.011	.032
23	.003	.014	.021	.000	.000	.004	.011	.023	.000	.000	.000	.033
24	.007	.014	.011	.022	.000	.008	.016	.012	.028	.000	.005	.000
25	.006	.000	.022	.000	.000	.008	.000	.024	.000	.000	.000	.000
26	.001	.029	.022	.000	.048	.002	.034	.025	.000	.057	.000	.000
27	.004	.010	.011	.000	.000	.006	.012	.000	.000	.000	.000	.000
28	.006	.005	.035	.022	.000	.004	.006	.013	.000	.000	.011	.000
29	.006	.015	.000	.000	.000	.006	.018	.000	.000	.000	.005	.000
30	.003	.005	.000	.047	.000	.004	.006	.000	.029	.000	.000	.000
31	.003	.010	.000	.000	.000	.004	.012	.000	.000	.000	.000	.000
32	.006	.000	.000	.000	.000	.004	.000	.000	.000	.000	.011	.000
33	.005	.016	.000	.000	.000	.004	.019	.000	.000	.000	.006	.000
34	.005	.000	.000	.000	.000	.006	.000	.000	.000	.000	.000	.000
35	.008	.005	.012	.025	.000	.008	.000	.000	.031	.000	.006	.035
36	.003	.011	.012	.000	.000	.004	.006	.013	.000	.000	.000	.036
37	.009	.005	.012	.000	.000	.011	.006	.013	.000	.000	.006	.000
38	.000	.005	.013	.000	.000	.000	.006	.014	.000	.000	.000	.000
39	.005	.022	.000	.000	.000	.006	.026	.000	.000	.000	.000	.000
40	.005	.000	.013	.000	.000	.007	.000	.014	.000	.000	.000	.000
41	.003	.006	.013	.000	.000	.004	.007	.014	.000	.000	.000	.000

Table 3.5 (cont.)

	Total					Blacks					Whites	
Month	1st (N = 911)	2nd (N = 325)	3rd (N = 157)	4th (N = 83)	5th (N = 40)	1st (N = 693)	2nd (N = 277)	3rd (N = 137)	4th (N = 69)	5th (N = 34)	1st (N = 218)	2nd (N = 48)
42	.005	.006	.000	.026	.000	.007	.007	.000	.033	.000	.000	.000
43	.008	.006	.000	.000	.000	.009	.007	.000	.000	.000	.006	.000
44	.002	.012	.013	.000	.000	.002	.014	.015	.000	.000	.000	.000
45	.003	.000	.000	.000	.055	.005	.000	.000	.000	.068	.000	.000
46	.010	.006	.014	.000	.067	.009	.007	.015	.000	.095	.011	.000
47	.002	.006	.029	.000	.000	.002	.007	.032	.000	.000	.000	.000
48	.005	.018	.015	.000	.000	.007	.022	.017	.000	.000	.000	.000
49	.005	.000	.000	.000	.000	.007	.000	.000	.000	.000	.000	.000
50	.003	.006	.000	.000	.000	.002	.007	.000	.000	.000	.006	.000
51	.000	.006	.000	.000	.000	.000	.008	.000	.000	.000	.000	.000
52	.003	.006	.016	.000	.000	.005	.008	.018	.000	.000	.000	.000
53	.003	.006	.016	.000	.000	.005	.008	.018	.000	.000	.000	.000
54	.002	.000	.000	.000	.000	.002	.000	.000	.000	.000	.000	.000
55	.002	.013	.000	.000	.000	.002	.016	.000	.000	.000	.000	.000
56	.007	.000	.000	.000	.000	.007	.000	.000	.000	.000	.006	.000
57	.002	.000	.000	.000	.000	.002	.000	.000	.000	.000	.000	.000
58	.002	.000	.000	.000	.000	.002	.000	.000	.000	.000	.000	.000
59	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
60	.002	.000	.000	.000	.000	.002	.000	.000	.000	.000	.000	.000
61	.005	.000	.000	.000	.000	.005	.000	.000	.000	.000	.006	.000
62	.002	.007	.000	.000	.000	.002	.008	.000	.000	.000	.000	.000
63	.002	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000	.000
64	.000	.014	.000	.000	.000	.000	.017	.000	.000	.000	.000	.000
65	.004	.000	.000	.000	.000	.005	.000	.000	.000	.000	.000	.000
66	.002	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000	.000
67	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
68	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
69	.002	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000	.000
70	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
71	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
72	.002	.007	.000	.000	.000	.003	.009	.000	.000	.000	.000	.000
73	.000	.000	.000	.000	.000	.000	.000	.000	NA[b]	.000	.000	.000
74	.002	.000	.000	.000	.000	.000	.000	.000	NA	.000	.006	.000
75	.000	.000	.000	.000	.000	.000	.000	.000	NA	.000	.000	.000
76	.004	.000	.000	.000	.000	.005	.000	.000	NA	.000	.000	.000
77	.004	.000	.000	.000	.000	.003	.000	.000	NA	.000	.006	.000
78	.002	.000	.017	.000	.000	.003	.000	.000	NA	.000	.000	.000
79	.000	.000	.000	.000	.000	.000	.000	.000	NA	.000	.000	.000
80	.000	.000	.000	.000	.000	.000	.000	.000	NA	.000	.000	.000
81	.000	.000	.000	.000	NA	.000	.000	.000	NA	NA	.000	.000
82	.002	.C00	.000	.000	NA	.000	.000	.000	NA	NA	.000	.000
83	.000	.000	.000	.000	NA	.003	.000	.000	NA	NA	.000	.000
84	.000	.000	.000	.000	NA	.000	.000	.000	NA	NA	.000	.000
85	.000	.000	.000	.000	NA	.000	.000	.000	NA	NA	.000	.000
86	.004	.000	.000	.000	NA	.005	.000	.000	NA	NA	.000	.000

Table 3.5 (cont.)

	Total					Blacks					Whites	
Month	1st (N = 911)	2nd (N = 325)	3rd (N = 157)	4th (N = 83)	5th (N = 40)	1st (N = 693)	2nd (N = 277)	3rd (N = 137)	4th (N = 69)	5th (N = 34)	1st (N = 218)	2nd (N = 48)
87	.000	.000	.000	.000	NA	.000	.000	.000	NA	NA	.000	.000
88	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000	.000
89	.002	.000	.000	NA	NA	.000	.000	.000	NA	NA	.006	.000
90	.002	.000	.000	NA	NA	.003	.000	.000	NA	NA	.000	.000
91	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000	.000
92	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000	.000
93	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000	.000
94	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000	.000
95	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000	.000
96	.000	.000	.000	NA	NA	.000	.000	.000	NA	NA	.000	.000

a. The number of birth cohort subjects at risk of rearrest.

b. The cell entry is not applicable because no birth cohort subjects were exposed to the risk of rearrest during this month.

Table 3.6

1945 Birth Cohort: Arrests for Violent Crimes--
Loglikelihood Statistic by Type of Parametric
Distribution, Age Status, Race, and Arrest
Transition

		Arrest Transition															
		Juveniles (Total Sample)										Adults (Follow-up Sample)					
		Total				Blacks				Whites		Total		Blacks		Whites	
		1st (N = 360)[a] (R = 75)[b]		2nd (N = 75) (R = 22)		1st (N = 302) (R = 72)		2nd (N = 72) (R = 20)		1st (N = 58) (R = 3)		1st (N = 74) (R = 30)		1st (N = 56) (R = 26)		1st (N = 18) (R = 4)	
Parametric Distribution	Number of Parameters[c]	LL[d]	% R[e]	LL	% R	LL	% R	LL	% R	LL	% R	LL	% R	LL	% R	LL	% R
Exponential	1	-443	100	-114	100	-416	100	-105	100	-22	100	-172	100	-145	100	-26	100
Split Exponential	2	-439	44	-110	44	-412	46	-102	44	--[f]	--	-168	50	-141	56	-25	29
Loglogistic	2	-438	100	-108	100	-411	100	-101	100	-21	100	-169	100	-142	100	-25	100
Split Loglogistic	3	--	--	-108	91	-411	86	--	--	--	--	-168	64	-141	77	-22	28
Lognormal	2	-439	100	-107	100	-412	100	-101	100	-22	100	-168	100	-141	100	-25	100
Split Lognormal	3	--	--	--	--	--	--	--	--	--	--	-168	75	-141	91	-23	27
Weibull	2	-438	100	-108	100	-411	100	-101	100	-21	100	-169	100	-142	100	-26	100
Split Weibull	3	-438	60	-108	60	-411	55	-101	65	--	--	-167	50	-141	57	-22	30
Gompertz	2	-439	50	-110	46	-412	51	-103	46	-22	100	-168	53	-142	59	-25	30
Mixed Exponential	3	-438	100	--	--	-411	100	--	--	--	--	--	--	--	--	--	--

a. The number of birth cohort subjects at risk of rearrest.

b. The number of rearrested birth cohort subjects.

c. The number of parameters characterizing the distribution.

d. The distribution loglikelihood statistic.

e. The percentage of birth cohort subjects estimated to be rearrested by the distribution.

f. The loglikelihood statistic and corresponding percentage rearrested were not reported for this split-population distribution because the results were identical to the related unitary-population distribution. Identical results can occur for the split forms of the exponential, loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that the findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.7

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Loglikelihood Statistic by Type of Parametric
Distribution, Race, and Arrest Transition
(Construction Sample)

		Arrest Transition									
		Total									
		1st (N = 759)[a] (R = 262)[b]		2nd (N = 262) (R = 124)		3rd (N = 124) (R = 62)		4th (N = 62) (R = 37)		5th (N = 37) (R = 23)	
Parametric Distribution	Number of Parameters[c]	LL[d]	% R[e]	LL	% R	LL	% R	LL	% R	LL	% R
Exponential	1	-1363	100	-532	100	-259	100	-141	100	-74	100
Split Exponential	2	-1351	62	-527	78	-250	68	-141	97	-71	81
Loglogistic	2	-1341	100	-526	100	-245	100	-140	100	-70	100
Split Loglogistic	3	--	--	--	--	-244	86	--	--	-70	92
Lognormal	2	-1340	100	-526	100	-242	100	-140	100	-70	100
Split Lognormal	3	--[f]	--	--	--	-242	87	--	--	-70	90
Weibull	2	-1341	100	-527	100	-246	100	-140	100	-71	100
Split Weibull	3	-1341	91	-526	83	-245	76	--	--	-70	82
Gompertz	2	-1350	68	-528	91	-250	73	-141	99	-72	88
Mixed Exponential	3	-1339	100	-526	100	-243	100	-139	100	--	--

Table 3.7 (cont.)

		Arrest Transition											
		Blacks										Whites	
		1st (N = 644)[a] (R = 245)[b]		2nd (N = 245) (R = 117)		3rd (N = 117) (R = 59)		4th (N = 59) (R = 36)		5th (N = 36) (R = 22)		1st (N = 115) (R = 17)	
Parametric Distribution	Number of Parameters[c]	LL[d]	% R[e]	LL	% R	LL	% R	LL	% R	LL	% R	LL	% R
Exponential	1	-1247	100	-500	100	-247	100	-138	100	-72	100	-104	100
Split Exponential	2	-1235	65	-496	79	-239	68	-137	96	-68	81	-104	50
Loglogistic	2	-1226	100	-496	100	-233	100	-136	100	-68	100	-104	100
Split Loglogistic	3	--	--	--	--	-233	86	--	--	-68	93	--	--
Lognormal	2	-1225	100	-495	100	-231	100	-136	100	-67	100	-104	100
Split Lognormal	3	--	--	--	--	-231	87	--	--	-67	90	--	--
Weibull	2	-1226	100	-496	100	-235	100	-136	100	-69	100	-104	100
Split Weibull	3	-1226	90	-495	85	-234	77	--	--	-68	82	--	--
Gompertz	2	-1234	72	-497	93	-238	73	-137	99	-69	88	-104	58
Mixed Exponential	3	-1224	100	-495	100	-231	100	-135	100	--	--	-103	100

a. The number of birth cohort subjects at risk of rearrest.

b. The number of rearrested birth cohort subjects.

c. The number of parameters characterizing the distribution.

d. The distribution loglikelihood statistic.

e. The percentage of birth cohort subjects estimated to be rearrested by the distribution.

f. The loglikelihood statistic and corresponding percentage rearrested were not reported for this split-population distribution because the results were identical to the related unitary-population distribution. Identical results can occur for the split forms of the exponential, loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that the findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.8

1958 Birth Cohort: Adult Arrests for Violent Crimes--
Loglikelihood Statistic by Type of Parametric
Distribution, Race, and Arrest Transition
(Construction Sample)

		Arrest Transition									
		Total									
		1st (N = 911)[a] (R = 325)[b]		2nd (N = 325) (R = 157)		3rd (N = 157) (R = 83)		4th (N = 83) (R = 40)		5th (N = 40) (R = 21)	
Parametric Distribution	Number of Parameters[c]	LL[d]	% R[e]	LL	% R	LL	% R	LL	% R	LL	% R
Exponential	1	-1922	100	-832	100	-418	100	-198	100	-99	100
Split Exponential	2	-1859	42	-805	58	-412	67	-188	56	-94	60
Loglogistic	2	-1847	100	-802	100	-412	100	-187	100	-92	100
Split Loglogistic	3	-1844	63	-801	82	-412	90	-187	77	-91	70
Lognormal	2	-1841	100	-800	100	-412	100	-186	100	-91	100
Split Lognormal	3	-1840	76	-800	92	--[f]	--	-186	80	-91	71
Weibull	2	-1852	100	-805	100	-413	100	-188	100	-93	100
Split Weibull	3	-1845	48	-801	63	-412	70	-187	60	-92	65
Gompertz	2	-1860	44	-807	60	-413	73	-189	58	-94	63
Mixed Exponential	3	-1849	100	--	--	-412	100	--	--	-91	100

Table 3.8 (cont.)

Parametric Distribution	Number of Parameters[c]	Arrest Transition: Blacks 1st (N = 693)[a] (R = 277)[b] LL[d]	% R[e]	Blacks 2nd (N = 277) (R = 137) LL	% R	Blacks 3rd (N = 137) (R = 69) LL	% R	Blacks 4th (N = 69) (R = 34) LL	% R	Blacks 5th (N = 34) (R = 18) LL	% R	Whites 1st (N = 218) (R = 48) LL	% R	Whites 2nd (N = 48) (R = 20) LL	% R
Exponential	1	-1590	100	-724	100	-351	100	-165	100	-83	100	-315	100	-108	100
Split Exponential	2	-1540	48	-705	61	-345	64	-157	58	-81	64	-305	26	- 96	44
Loglogistic	2	-1530	100	-704	100	-345	100	-157	100	-78	100	-303	100	- 95	100
Split Loglogistic	3	-1528	71	-704	84	-344	88	-157	78	-78	84	-303	41	- 93	49
Lognormal	2	-1525	100	-703	100	-344	100	-157	100	-78	100	-302	100	- 94	100
Split Lognormal	3	-1525	86	-703	98	--	--	-157	82	-78	87	-302	49	- 92	48
Weibull	2	-1534	100	-707	100	-346	100	-158	100	-79	100	-304	100	- 96	100
Split Weibull	3	-1528	54	-703	64	-344	67	-157	61	-79	77	-303	30	- 93	45
Gompertz	2	-1541	49	-706	64	-345	70	-158	60	-80	67	-305	27	- 96	45
Mixed Exponential	3	-1531	100	--	--	--	--	--	--	-78	100	-303	100	- 93	100

a. The number of birth cohort subjects at risk of rearrest.

b. The number of rearrested birth cohort subjects.

c. The number of parameters characterizing the distribution.

d. The distribution loglikelihood statistic.

e. The percentage of birth cohort subjects estimated to be rearrested by the distribution.

f. The loglikelihood statistic and corresponding percentage rearrested were not reported for this split-population distribution because the results were identical to the related unitary-population distribution. Identical results can occur for the split forms of the exponential, loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that the findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.9

1945 Birth Cohort: Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles
by Type of Distribution, Age Status, Race,
and Arrest Transition

	Juveniles (Total Sample)																			
	Total								Blacks								Whites			
	1st				2nd				1st				2nd				1st			
Distribution	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90
Observed	12	35	NA[a]	NA	1	16	NA	NA	11	31	NA	NA	2	18	NA	NA	70	71	NA	NA
Exponential	14	39	94	312	7	19	45	151	13	34	83	275	8	21	50	165	55	149	359	1,192
Split Exponential	11	36	NA	NA	4	13	NA	NA	9	30	NA	NA	5	15	NA	NA	--[b]	--	--	--
Loglogistic	10	38	143	1,995	2	14	79	2,561	9	32	118	1,592	3	17	87	2,408	67	311	1,444	88,587
Split Loglogistic	--	--	--	--	2	14	84	78,441	9	32	128	NA	--	--	--	--	--	--	--	--
Lognormal	9	39	198	4,325	2	13	91	3,707	8	32	153	3,035	3	16	105	3,945	81	720	8,109	2,986,361
Split Lognormal	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Weibull	11	39	123	589	3	15	69	569	9	33	104	501	3	17	75	552	65	265	914	7,175
Split Weibull	10	38	162	NA	3	14	90	NA	9	31	157	NA	3	16	89	NA	--	--	--	--
Gompertz	11	36	NA	NA	4	13	NA	NA	10	31	202	NA	5	15	NA	NA	53	114	194	367
Mixed Exponential	11	37	160	772	--	--	--	--	9	30	157	1,002	--	--	--	--	--	--	--	--

Table 3.9 (cont.)

	Adults (Follow-up Sample)											
	Total				Blacks				Whites			
	1st				1st				1st			
Distribution	10	25	50	90	10	25	50	90	10	25	50	90
Observed	8	23	NA	NA	4	20	55	NA	22	26	NA	NA
Exponential	12	33	79	263	10	28	67	222	24	66	159	528
Split Exponential	7	21	152	NA	6	18	67	NA	12	57	NA	NA
Loglogistic	6	24	89	1,226	5	19	70	991	19	61	189	1,832
Split Loglogistic	7	22	107	NA	5	18	72	NA	19	36	NA	NA
Lognormal	6	23	96	1,493	5	17	73	1,129	21	60	196	1,843
Split Lognormal	6	21	104	NA	5	17	73	7,641	18	37	NA	NA
Weibull	6	26	89	484	5	20	71	401	20	66	186	768
Split Weibull	7	21	NA	NA	6	18	68	NA	18	29	NA	NA
Gompertz	7	21	116	NA	6	18	69	NA	13	62	NA	NA
Mixed Exponential	--	--	--	--	--	--	--	--	--	--	--	--

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

b. The rearrest percentile was not reported for this split population distribution because the results were identical to those for the related unitary-population distribution. Identical results can occur for the split forms of the exponential loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that these findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.10

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles by Type
of Distribution and Arrest Transition
(Construction Sample)

	Arrest Transition																			
	1st				2nd				3rd				4th				5th			
Distribution	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90
Observed	4	15	57	NA[a]	2	7	18	NA	1	3	17	NA	1	4	10	45	<.5	1	5	NA
Exponential	7	19	46	154	3	8	19	62	3	7	17	55	2	5	12	39	1	3	6	21
Split Exponential	6	16	52	NA	2	7	17	NA	2	5	14	NA	2	5	11	42	1	2	5	NA
Loglogistic	4	15	56	773	2	6	18	157	1	4	14	237	1	3	10	89	<.5	2	5	42
Split Loglogistic	--[b]	--	--	--	--	--	--	--	1	3	14	NA	--	--	--	--	<.5	2	5	111
Lognormal	4	14	62	1,016	2	6	18	171	1	3	15	224	1	3	10	87	1	2	5	39
Split Lognormal	--	--	--	--	--	--	--	--	1	3	14	NA	--	--	--	--	1	2	4	NA
Weibull	4	16	53	269	2	7	18	77	1	4	16	105	1	4	11	48	<.5	2	6	28
Split Weibull	4	16	53	570	2	6	18	NA	1	4	15	NA	--	--	--	--	1	2	5	NA
Gompertz	5	16	51	NA	2	7	18	167	2	4	13	NA	2	4	11	45	1	2	5	NA
Mixed Exponential	4	16	54	213	2	6	18	80	1	3	18	86	1	4	12	43	--	--	--	--

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

b. The rearrest percentile was not reported for this split population distribution because the results were identical to those for the related unitary-population distribution. Identical results can occur for the split forms of the exponential loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that these findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.11

1958 Birth Cohort: Black Juvenile Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles by Type of
Distribution and Arrest Transition
(Construction Sample)

	Arrest Transition																			
	1st				2nd				3rd				4th				5th			
Distribution	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90
Observed	3	13	48	NA[a]	2	7	18	NA	1	3	17	NA	1	4	10	45	<.5	1	6	NA
Exponential	6	17	41	138	3	8	18	61	3	7	17	56	2	5	12	39	1	3	7	22
Split Exponential	5	14	43	NA	2	7	17	NA	2	5	14	NA	2	5	11	43	1	2	5	NA
Loglogistic	4	13	47	629	2	6	18	148	1	4	14	241	1	3	10	94	1	2	5	47
Split Loglogistic	--[b]	--	--	--	--	--	--	--	1	3	14	NA	--	--	--	--	1	2	5	104
Lognormal	3	12	51	783	2	6	18	162	1	3	15	226	1	3	10	90	1	2	5	43
Split Lognormal	--	--	--	--	--	--	--	--	1	3	14	NA	--	--	--	--	1	2	5	183
Weibull	4	14	46	232	2	7	18	74	1	4	16	107	1	4	11	49	<.5	2	6	30
Split Weibull	4	14	46	864	2	6	18	NA	1	4	15	NA	--	--	--	--	1	2	5	NA
Gompertz	5	14	43	NA	2	7	18	113	2	4	13	NA	2	4	11	47	1	2	5	NA
Mixed Exponential	3	14	47	180	2	6	18	79	1	3	18	87	1	4	12	44	--	--	--	--

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

b. The rearrest percentile was not reported for this split population distribution because the results were identical to those for the related unitary-population distribution. Identical results can occur for the split forms of the exponential loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that these findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.12

1958 Birth Cohort: White Juvenile Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles by Type of
Distribution and Arrest Transition
(Construction Sample)

	Arrest Transition			
	1st			
Distribution	10	25	50	90
Observed	12	64	71	NA[a]
Exponential	18	49	119	394
Split Exponential	16	51	NA	NA
Loglogistic	15	55	202	2,763
Split Loglogistic	--[b]	--	--	--
Lognormal	14	59	305	6,888
Split Lognormal	--	--	--	--
Weibull	15	53	160	721
Split Weibull	--	--	--	--
Gompertz	16	50	200	NA
Mixed Exponential	13	58	166	597

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

b. The rearrest percentile was not reported for this split population distribution because the results were identical to those for the related unitary-population distribution. Identical results can occur for the split forms of the exponential loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that these findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.13

1958 Birth Cohort: Adults Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles by
Type of Distribution and Arrest Transition
(Construction Sample)

	Arrest Transition																			
	1st				2nd				3rd				4th				5th			
Distribution	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90
Observed	5	27	NA[a]	NA	2	12	48	NA	4	12	39	NA	1	11	30	NA	1	4	26	NA
Exponential	14	39	95	314	8	21	51	170	6	16	40	131	5	15	36	119	4	12	28	93
Split Exponential	8	25	NA	NA	4	13	47	NA	4	12	36	NA	3	8	31	NA	2	7	22	NA
Loglogistic	5	27	141	3,863	3	12	52	964	4	12	38	397	2	8	36	815	1	5	24	548
Split Loglogistic	5	24	220	NA	3	12	53	NA	4	11	37	130,000	2	7	37	NA	1	4	23	NA
Lognormal	5	26	159	5,107	3	11	54	1,046	3	11	39	444	2	7	37	856	1	5	25	512
Split Lognormal	5	24	187	NA	3	11	54	3,329	--[b]	--	--	--	2	7	38	NA	1	4	24	NA
Weibull	5	29	130	1,003	3	13	54	361	3	12	39	190	2	8	38	290	1	6	27	223
Split Weibull	5	25	NA	NA	3	12	51	NA	4	12	36	NA	2	7	35	NA	1	5	24	NA
Gompertz	7	25	NA	NA	4	13	48	NA	4	12	36	NA	3	8	32	NA	2	6	21	NA
Mixed Exponential	5	27	124	512	--	--	--	--	4	12	36	585	--	--	--	--	1	4	27	NA

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

b. The rearrest percentile was not reported for this split population distribution because the results were identical to those for the related unitary-population distribution. Identical results can occur for the split forms of the exponential loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that these findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.14

1958 Birth Cohort: Black Adult Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles by Type
of Distribution and Rearrest Transition
(Construction Sample)

	Arrest Transition																			
	1st				2nd				3rd				4th				5th			
Distribution	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90
Observed	4	20	NA[a]	NA	3	13	47	NA	4	12	43	NA	1	10	30	NA	1	4	26	NA
Exponential	12	33	79	264	8	21	50	167	6	17	41	136	5	13	32	107	4	11	26	87
Split Exponential	7	21	NA	NA	5	14	45	NA	4	12	38	NA	3	8	28	NA	3	7	22	NA
Loglogistic	4	21	103	2,496	3	13	50	746	4	12	40	462	2	7	32	587	1	5	23	506
Split Loglogistic	4	19	122	NA	3	12	50	NA	4	12	40	NA	2	7	32	NA	1	5	23	NA
Lognormal	4	20	111	3,048	3	12	51	814	3	11	42	519	2	7	33	622	1	5	23	470
Split Lognormal	4	19	117	NA	3	12	51	978	--[b]	--	--	--	2	7	33	NA	1	5	23	NA
Weibull	4	23	98	729	3	14	52	307	3	13	42	209	2	8	33	229	1	6	25	196
Split Weibull	4	20	133	NA	4	13	48	NA	4	12	39	NA	2	7	30	NA	1	5	24	NA
Gompertz	6	20	NA	NA	4	13	46	NA	4	12	39	NA	3	8	29	NA	2	6	21	NA
Mixed Exponential	4	20	98	416	--	--	--	--	--	--	--	--	--	--	--	--	1	4	27	161

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

b. The rearrest percentile was not reported for this split population distribution because the results were identical to those for the related unitary-population distribution. Identical results can occur for the split forms of the exponential loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that these findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.15

1958 Birth Cohort: White Adult Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles by Type of
Distribution and Arrest Transition
(Construction Sample)

	Arrest Transition							
	1st				2nd			
Distribution	10	25	50	90	10	25	50	90
Observed	12	88	NA[a]	NA	1	6	NA	NA
Exponential	28	75	181	603	9	23	56	187
Split Exponential	14	92	NA	NA	3	9	NA	NA
Loglogistic	13	82	450	19,000	1	9	75	5,618
Split Loglogistic	12	89	NA	NA	1	6	NA	NA
Lognormal	12	84	705	40,000	1	9	80	5,318
Split Lognormal	11	89	NA	NA	1	6	NA	NA
Weibull	14	82	391	3,276	1	10	78	1,252
Split Weibull	12	89	NA	NA	1	7	NA	NA
Gompertz	14	87	NA	NA	2	8	NA	NA
Mixed Exponential	12	88	337	1,329	1	5	82	413

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

Table 3.16

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Failure-Time Regression-Model Loglikelihoods, Shape and
Scale Parameters, and Significant Risk Variables by
Arrest Transition, Race, Parametric Distribution,
and Risk Variable Model
(Construction Sample)

Panel A: 1st Arrest Transition--Total, Blacks, Whites
Panel B: 2nd Arrest Transition--Blacks
Panel C: 3rd Arrest Transition--Blacks
Panel D: 4th Arrest Transition--Blacks
Panel E: 5th Arrest Transition--Blacks

Panel A.1: 1st Arrest Transition--Total (N = 759)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA[k]	4	5	22	0	4	5	22	0	4	5	22	0	4	5	22	0	4	5	22
-2 Loglikelihood[e]	NA	3,138	3,120*	3,052*	4,510	4,502	4,480*	4,394*	4,446	4,440	4,420*	4,356*	4,454	4,448	4,430*	4,366*	4,476	4,472	4,454*	4,402*
Shape[f]	NA	--	--	--	1.0	1.0	1.0	1.0	1.5*	1.5*	1.5*	1.4*	1.3*	1.3*	1.3*	1.2*	2.6	2.6	2.6	2.5
Scale[f]	NA	--	--	--	7.6	8.3	9.0	11.7	8.0	8.9	9.9	11.7	7.6	8.2	9.4	11.2	7.8	8.5	9.4	11.2
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NS	NS	NA	- .4	- .3	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
.Weapon Used																				
-Firearm	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Other Weapon	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	1.0	.9	NA	NA	-1.0	-.9	NA	NA	-1.4	-1.3	NA	NA	-1.5	-1.3	NA	NA	-1.6	-1.5
.Prior Status Offense	NA	NA	NA	.6	NA	NA	NA	-.6	NA	NA	NA	- .8	NA	NA	NA	- .8	NA	NA	NA	- .9
.Age at Arrest for Present Violent Crime	NA	NA	NA	<.1	NA	NA	NA	>-.1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
.Age at First Arrest	NA	NA	NA	NS	NA	NA	NA	< .1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS

Table 3.16 (cont.)

Panel A.2: 1st Arrest Transition--Blacks (N = 644)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21
-2 Loglikelihood[e]	NA	2,858	NA	2,798*	4,162	4,156	NA	4,076*	4,104	4,100	NA	4,042*	4,110	4,108	NA	4,052*	4,132	4,130	NA	4,082*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.5*	1.5*	NA	1.4*	1.3*	1.3*	NA	1.2*	2.6	2.6	NA	2.4
Scale[f]	NA	--	NA	--	7.5	8.0	NA	11.0	7.8	8.4	NA	10.9	7.4	7.8	NA	10.5*	7.5	7.9	NA	10.1
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
.Weapon Used																				
-Firearm	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Other Weapon	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	NA	NS	NS	NS
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Prior Status Offense	NA	NA	NA	.6	NA	NA	NA	-.7	NA	NA	NA	-.9	NA	NA	NA	-.9	NA	NA	NA	-1.0
.Age at Arrest for Present Violent Crime	NA	NA	NA	NA	NA	NA	NA	>-.1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS

Table 3.16 (cont.)

Panel A.3: 1st Arrest Transition--Whites (N = 115)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21
-2 Loglikelihood[e]	NA	130	NA	88	324	318	NA	276	320	314	NA	286·	322	314	NA	291	324	316	NA	300
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.5*	1.5*	NA	NC	1.4	1.4	NA	NC	3.1	3.0	NA	NC
Scale[f]	NA	--	NA	--	8.5	9.0	NA	NC	9.3	10.0	NA	NC	9.1	10.0	NA	NC	9.9	10.5	NA	NC
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC
.Weapon Used																				
-Firearm	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC
-Other Weapon	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.16 (cont.)

Panel B: 2nd Arrest Transition--Blacks (N = 245)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25
-2 Loglikelihood[e]	NA	1,123	NA	1,090*	1,796	1,788	NA	1,755*	1,786	1,780	NA	1,750	1,790	1,784	NA	1,758	1,802	1,798	NA	1,770
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.3*	1.2*	NA	1.2*	1.1	1.0	NA	1.0	2.1	2.0	NA	2.0
Scale[f]	NA	--	NA	--	6.7	6.2	NA	4.1	6.8	6.1	NA	3.1	6.3	5.2	NA	1.9	6.4	5.2	NA	1.0
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	-.2	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	.6	NA	.8	NA	-.7	NA	-.8	NA	NS	NA	- .9	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NS	NA	NA	NA	NA	NA	NA	NA	NA
.Age at Arrest for Present Violent Crime	NA	NA	NA	>-.1	NA	NA	NA	< .1	NA	NA	NA	< .1	NA	NA	NA	< .1	NA	NA	NA	< .1
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Any Priors																				
-Yes	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown	NA	NA	NA	3.1	NA	NA	NA	-3.4	NA	NA	NA	-3.7	NA	NA	NA	-3.3	NA	NA	NA	-3.8
-No [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Mean Seriousness																				
-Known Adjudicated/Convicted	NA	NA	NA	- .5	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown Adjudicated/Convicted	NA	NA	NA	-1.0	NA	NA	NA	NS	NA	NA	NA	1.2	NA	NA	NA	1.1	NA	NA	NA	1.2
.Socioeconomic Status ≤ 15th Percentile	NA	NA	NA	NS	NA	NA	NA	.4	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS

Table 3.16 (cont.)

Panel C: 3rd Arrest Transition--Blacks (N = 117)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25
-2 Loglikelihood[e]	NA	486	NA	468	896	890	NA	856*	860	856	NA	838	860	856	NA	840	860	856	NA	840
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.8*	1.8*	NA	1.6*	1.5*	1.4*	NA	1.3*	2.6	2.6	NA	2.4
Scale[f]	NA	--	NA	--	6.6	5.7	NA	6.1	6.8	5.7	NA	3.6	6.1	5.5	NA	2.4	6.2	5.7	NA	2.4
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	.7	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Prior Arrests for UCR Index Crimes																				
.Mean Seriousness	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	3.2	NA	NA	NA	3.6	NA	NA	NA	3.7
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Any Priors																				
-Yes	NA	NA	NA	NS	NA	NA	NA	-3.3	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown	NA	NS	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-No [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Mean Seriousness -Unknown Adjudicated/Convicted	NA	NA	NA	-1.9	NA	NA	NA	2.4	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
Most Recent Prior UCR Index Crime																				
.Seriousness (Log)	NA	NA	NA	NS	NA	NA	NA	- .7	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS

Table 3.16 (cont.)

Panel D: 4th Arrest Transition--Blacks (N = 59)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25
-2 Loglikelihood[e]	NA	230	NA	204	520	512	NA	500	512	506	NA	498	516	508	NA	480	518	508*	NA	480
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.4*	1.4*	NA	NC	1.2	1.1	NA	.8	2.2	2.0	NA	1.4
Scale[f]	NA	--	NA	--	6.2	6.4	NA	NC[e]	6.3	6.5	NA	NC	5.8	6.3	NA	-7.0	5.7	6.3	NA	-4.2
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[e]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NS	NA	NS	NA	1.5
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	NA	NA	NA	NA	NA
.Age at Arrest for Present Violent Crime	NA	NA	NA	-.1	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	.1	NA	NA	NA	.1
Adjudicated/Convicted for Prior UCR Index Crimes																				
.Mean Seriousness -Known Adjudicated/Convicted	NA	NA	NA	2.2	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	NS	NA	NA	NA	NS
.Incarcerated for a Prior UCR Index Crime	NA	NA	NA	NS	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	1.3	NA	NA	NA	NS

Table 3.16--Panel D (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
First Prior UCR Index Crime																				
.Type																				
-Robbery	NA	NA	NA	-1.4	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	NS	NA	NA	NA	NS
-Assault	NA	NA	NA	NS	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	NS	NA	NA	NA	NS
-Property [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Most Recent Prior UCR Index Crime																				
.Type																				
-Robbery	NA	NA	NA	NS	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	NS	NA	NA	NA	NS
-Assault [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	1.5	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	-1.4	NA	NA	NA	NS
-Other Weapon	NA	NA	NA	1.2	NA	NA	NA	NC	NA	NA	NA	NC	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.16 (cont.)

Panel E: 5th Arrest Transition--Blacks (N = 36)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	0[c]	L[c]	L+R[c]	A[c]	0	L	L+R	A	0	L	L+R	A	0	L	L+R	A	0	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	NE[j]	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE
-2 Loglikelihood[e]	NA	124	NA	NE	292	288	NA	NE	286	282	NA	NE	286	280	NA	NE	286	280	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	1.5*	1.4*	NA	NE	1.1	1.0	NA	NE	2.0	1.8	NA	NE
Scale[f]	NA	--	NA	NE	5.6	9.1	NA	NE	5.7	9.8	NA	NE	5.0	8.9	NA	NE	5.0	9.2	NA	NE
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	-1.8	NA	NE	NA	-1.8	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NE	NA	-1.0	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

a. At each arrest transition, a proportional hazards (i.e., Cox) regression model was estimated in addition to the four parametric failure time regression models. This model permits one to estimate the effects of risk variables on the time-specific rearrest risk (i.e., the hazard rate) but, because it makes no assumption about the shape of the underlying parametric distribution which generated these risks, does not permit one to estimate distributional parameters. (Thus, it is a semiparametric model--parametric only in the coefficients.) The proportional hazards model assumes that the hazard functions of different levels, or strata, of a risk variable (e.g., low SES and high SES) are proportional to one another across time. For instance, low SES subjects might have a hazard rate that is twice as high as high SES subjects regardless of the point in time at which one contrasts the groups. While the proportional hazards model might have limited use with respect to js policy making which aims to devise interventions which are linked to time, because the model does not assign a specific time-varying shape to the hazard function, it does have an important feature: it is robust. Risk-variable coefficients are reliably estimated by the proportional hazards model across a variety of parametric distributions (e.g., exponential, Weibull, loglogistic). Because it is robust, the proportional hazards model can be used as a baseline model against which to compare the coefficient estimates produced by the other parametric models. We use the proportional hazards model, therefore, to check the consistency and, in turn, plausibility of our results.

As the name hints, the risk variables in the proportional hazards model influence the <u>hazard function</u>, which is the rearrest function in this study. A positive coefficient indicates that the presence of a risk characteristic increases the rearrest function; a negative coefficient indicates the reverse. To compare the effect of a coefficient estimated by the proportional hazards model to the effect of a coefficient estimated by one of the other parametric models, one simply flips the sign of the proportional hazards coefficient. This is done because the parametric models produce coefficient estimates based on the relationship of the risk variable to the timing of rearrest, not to the rearrest (i.e., hazard) rate. When a coefficient in the proportional

hazards model has a positive sign, the risk variable increases the rearrest risk, which corresponds to a negative sign in the parametric models, indicating a more rapid time until rearrest. Put somewhat differently, a higher hazard rate implies a shorter rearrest time, indicated by the opposite signs of the same coefficient in the proportional hazards and the parametric models.

To estimate a proportional hazards model, at least one risk variable must be included. For this reason, the unconditional risk-variable model (i.e., no risk variables included), designated by the tabular column "0," does not apply and is, therefore, not filled in. Also, because the proportional hazards model is semiparametric in the way described above, the tabular rows designating the "shape" and "scale" parameters do not apply and, likewise, are not filled in. See notes c and f for further discussion of the different risk-variable models which were estimated and of some special features of the shape and scale parameters.

b. We estimated the extreme value parameterization of this model. For details of the computational procedure, see D. Steinberg and P. Colla, SURVIVAL: A Supplementary Module for SYSTAT (Evanston, IL: SYSTAT, Inc, 1988). The extreme value parameterization yields coefficient estimates which are identical in magnitude and sign to those estimated by the nonextreme value parameterizations. Only the shape and scale parameters differ across parameterizations, and these are easily converted to one another. Technical estimation issues mainly influenced the decision to use this parameterization.

c. 0: The unconditional model (i.e., no risk variables included);
L: The legally-permissible risk-variable model;
L+R: The legally-permissible-plus-race risk-variable model;
A: The all (i.e., full) risk-variable model.

d. The number of risk variables in the model. This number sometimes changed across arrest transitions. A risk variable was included in the model at a particular arrest transition based on its distributional features. First, and most obviously, risk variables were included only when they could produce reliable estimates. For this reason, for example, the birth cohort subject's race during the juvenile period was included only at the first arrest transition; too few whites appeared at the later arrest transitions. Second, and less obviously, some variables were included in the model because of the way those variables were technically defined. For instance, the risk variable indicating the presence of an arrest for a prior UCR index crime was included only at the first arrest transition because the variable could take on different values only at this arrest transition. (The birth cohort subject might or might not have been previously arrested for a UCR property index crime.) At all subsequent transitions, the birth cohort subject must have had a prior arrest for a UCR index crime, at a minimum, the first arrest for a serious violent crime, which placed the subject in the study sample.

e. This value is minus two times the model's loglikelihood statistic, a statistic measuring how well the model matched the observed rearrest-time data. We have calculated minus two times this statistic because this value can be used to judge the comparative merits of nested risk-variable models, that is, of models whose coefficients are related in the form of superset to subset. One can compare, then, the "0" model to the "L," "L+R," and "A" models, the "L" model to the "L+R" and "A" models, and the "L+R" to the "A" model. These comparisons are transitive: if the "L" model is significantly superior to the "0" model, the "L+R" and the "A" models are also statistically superior because they include the "L" model.

When statistically comparing two risk-variable models, we performed the following steps: (1) calculated the difference between -2 times the loglikelihood of each model, which we have presented in the table, (2) calculated the degrees of freedom of the statistical test, which is the difference between the numbers of risk variables in the two models, (3) turned to a chi-square table, found the tabular cell entry which was at the intersection of the degrees-of-freedom and selected significance-level, and determined whether the value calculated in step 1 was greater than the tabular cell entry, and (4) reported that the comparison was statistically significant if the value calculated in step 1 was greater than the cell entry or, conversely, that the comparison was statistically nonsignificant if the opposite was true. An asterisk ("*") appearing after a loglikelihood value indicates that the broader risk-variable model (i.e., the superset) significantly improved (p. val. < .05) the explanatory capacity relative to the narrower risk-variable model (i.e., the subset) immediately adjacent to the left.

In some instances, the risk-variable model failed to converge. In such cases, it was impermissible to draw formal inferences about the magnitudes and signs of the model's coefficients. However, one can still loosely use the loglikelihood statistic of a nonconvergent model to compare the explanatory capacity of that model to other, related risk-variable models. We did not, however, formally compare a risk-variable model which failed to converge to any other model.

NC: The model did not converge.

f. The distribution's shape parameter defines the curvature of the hazard function. For some distributions, the shape parameter invariably produces a specific type of overall curvature, although the shape parameter's magnitude effects the details of that curvature. (For instance, the lognormal distribution is always single peaked, but the shape parameter governs the rate of incline to the peak and, in turn, the rate of decline thereafter.) For

Table 3.16 (cont.)

other distributions, the shape parameter produces generically different curvatures depending upon the magnitude of the parameter. The overall curvatures of the hazard functions of the Weibull and loglogistic distributions depend upon the magnitudes their shape parameters as follows:

Distribution	Value of Shape Parameter
Weibull (extreme value)	= 1, constant hazard function (i.e., the Weibull reduces to the exponential form), > 1, decreasing hazard function, < 1, increasing hazard function.
loglogistic	> 1, decreasing hazard function, ≤ 1, single-peaked hazard function.

An asterisk ("*") after the shape parameter indicates that the parameter was significantly different (p. val. < .05) from 1, in the indicated direction.

g. The "bullet" (".") before the variable name indicates the general name of the risk variable; the dash before the variable name indicates a specific level of the general risk variable.

h. [REF]: The omitted reference category of the categorical variable. The effect of a specific category of a categorical variable (e.g., the presence of a firearm) on the timing of rearrest is obtained by comparing the coefficient calculated for that category to the reference category (e.g., the absence of a weapon).

i. Only those coefficients are presented which were significant at p. val. < .05.

j. NE: The model was not estimated because there were too few cases to produce reliable results.

k. NS: Not significant at p. val. < .05.

l. NA: Not applicable. The risk variable or risk-variable model did not apply at the arrest transition.

Table 3.17

1945 Birth Cohort: Juvenile Arrests for Violent Crimes--
Failure-Time Regression-Model Loglikelihoods, Shape and
Scale Parameters, and Significant Risk Variables by
Arrest Transition, Race, Parametric Distribution,
and Risk Variable Model
(Total Sample)

Panel A: 1st Arrest Transition--Total, Blacks, Whites
Panel B: 2nd Arrest Transition--Blacks

Panel A.1: 1st Arrest Transition--Total (N = 360)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21
-2 Loglikelihood[e]	NA	788	778*	752*	1,396	1,392	1,382*	1,350*	1,378	1,374	1,364*	1,342	1,380	1,376	1,364*	1,344	1,392	1,384	1,374*	1,352
Shape[f]	NA	--	--	--	1.0	1.0	1.0	1.0	1.5*	1.5*	1.5*	1.4*	1.4*	1.4*	1.4*	1.3	3.0	3.0	2.9	2.7
Scale[f]	NA	--	--	--	8.3	8.6	9.6	15.4	9.0	9.3	10.9	16.5	8.6	8.9	10.5	16.8	9.3	9.2	11.0	18.2
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS[k]	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
.Weapon Used																				
-Firearm	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Other Weapon	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	1.5[l]	1.5	NA	NA	-1.5	-1.5	NA	NA	-2.3	-2.0	NA	NA	-2.3	-2.0	NA	NA	-2.5	-2.1
.Age at Arrest for Present Violent Crime	NA	NA	NA	> .1	NA	NA	NA	>- .1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
.Age at First Arrest	NA	NA	NA	>-.1	NA	NA	NA	< .1	NA	NA	NA	< .1	NA	NA	NA	< .1	NA	NA	NA	< .1

Table 3.17--Panel A.1 (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
Most Recent Prior UCR Index Crime																				
.Seriousness (Log)	NA	NA	NA	1.5	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-2.2
-Other Weapon	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.17 (cont.)

Panel A.2: 1st Arrest Transition--Blacks (N = 302)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21
-2 Loglikelihood[e]	NA	734	NA	708*	1,322	1,318	NA	1,286*	1,304	1,300	NA	1,273*	1,306	1,300	NA	1,280	1,316	1,308	NA	1,288
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.5*	1.5*	NA	1.4*	1.4*	1.4*	NA	1.3	3.0	2.9	NA	2.7
Scale[f]	NA	--	NA	--	8.2	8.0	NA	14.3	8.8	8.6	NA	15.0	8.4	8.2	NA	15.6	9.0	8.3	NA	17.7
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	-.9	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	-1.7	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	.3	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	.7	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Prior Status Offense	NA	NA	NA	.6	NA	NA	NA	-.6	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
.Age at Arrest for Present Violent Crime	NA	NA	NA	< .1	NA	NA	NA	>-.1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
.Age at First Arrest	NA	NA	NA	>-.1	NA	NA	NA	< .1	NA	NA	NA	<.1	NA	NA	NA	<.1	NA	NA	NA	< .1
Most Recent Prior UCR Index Crime																				
.Seriousness (Log)	NA	NA	NA	1.5	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-2.2
-Other Weapon	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	.4
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.17 (cont.)

Panel A.3: 1st Arrest Transition--Whites (N = 58)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21
-2 Loglikelihood[e]	NA	16	NA	NE	62	62	NA	NE	64	62	NA	NE	64	62	NA	NE	64	62	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	1.3*	NC	NA	NE	1.3*	NC	NA	NE	3.2	NC	NA	NE
Scale[f]	NA	--	NA	NE	9.7	NC	NA	NE	10.4	NC	NA	NE	10.4	NC	NA	NE	11.8	NC	NA	NE
I. <u>Permissible</u>																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE	NA	NC	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. <u>Less Permissible and Impermissible</u>																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.17 (cont.)

Panel B: 2nd Arrest Transition--Blacks (N = 72)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21
-2 Loglikelihood[e]	NA	146	NA	NE	346	342	NA	NE	322	320	NA	NE	324	320	NA	NE	324	320	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	2.3*	2.3*	NA	NE	2.2*	2.1*	NA	NE	4.2	4.2	NA	NE
Scale[f]	NA	--	NA	NE	7.7	8.2	NA	NE	9.0	10.0	NA	NE	8.5	9.2	NA	NE	9.0	10.0	NA	NE
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	-1.1	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

a. At each arrest transition, a proportional hazards (i.e., Cox) regression model was estimated in addition to the four parametric failure time regression models. This model permits one to estimate the effects of risk variables on the time-specific rearrest risk (i.e., the hazard rate) but, because it makes no assumption about the shape of the underlying parametric distribution which generated these risks, does not permit one to estimate distributional parameters. (Thus, it is a semiparametric model--parametric only in the coefficients.) The proportional hazards model assumes that the hazard functions of different levels, or strata, of a risk variable (e.g., low SES and high SES) are proportional to one another across time. For instance, low SES subjects might have a hazard rate that is twice as high as high SES subjects regardless of the point in time at which one contrasts the groups. While the proportional hazards model might have limited use with respect to js policy making which aims to devise interventions which are linked to time, because the model does not assign a specific time-varying shape to the hazard function, it does have an important feature: it is robust. Risk-variable coefficients are reliably estimated by the proportional hazards model across a variety of parametric distributions (e.g., exponential, Weibull, loglogistic). Because it is robust, the proportional hazards model can be used as a baseline model against which to compare the coefficient estimates produced by the other parametric models. We use the proportional hazards model, therefore, to check the consistency and, in turn, plausibility of our results.

As the name hints, the risk variables in the proportional hazards model influence the hazard function, which is the rearrest function in this study. A positive coefficient indicates that the presence of a risk characteristic increases the rearrest function; a negative coefficient indicates the reverse. To compare the effect of a coefficient estimated by the proportional hazards model to the effect of a coefficient estimated by one of the other parametric models, one simply flips the sign of the proportional hazards coefficient. This is done because the parametric models produce coefficient estimates based on the relationship of the risk variable to the timing of rearrest, not to the rearrest (i.e., hazard) rate. When a coefficient in the proportional

Table 3.17 (cont.)

hazards model has a positive sign, the risk variable increases the rearrest risk, which corresponds to a negative sign in the parametric models, indicating a more rapid time until rearrest. Put somewhat differently, a higher hazard rate implies a shorter rearrest time, indicated by the opposite signs of the same coefficient in the proportional hazards and the parametric models.

To estimate a proportional hazards model, at least one risk variable must be included. For this reason, the unconditional risk-variable model (i.e., no risk variables included), designated by the tabular column "0," does not apply and is, therefore, not filled in. Also, because the proportional hazards model is semiparametric in the way described above, the tabular rows designating the "shape" and "scale" parameters do not apply and, likewise, are not filled in. See notes c and f for further discussion of the different risk-variable models which were estimated and of some special features of the shape and scale parameters.

b. We estimated the extreme value parameterization of this model. For details of the computational procedure, see D. Steinberg and P. Colla, SURVIVAL: A Supplementary Module for SYSTAT (Evanston, IL: SYSTAT, Inc, 1988). The extreme value parameterization yields coefficient estimates which are identical in magnitude and sign to those estimated by the nonextreme value parameterizations. Only the shape and scale parameters differ across parameterizations, and these are easily converted to one another. Technical estimation issues mainly influenced the decision to use this parameterization.

c. 0: The unconditional model (i.e., no risk variables included);
L: The legally-permissible risk-variable model;
L+R: The legally-permissible-plus-race risk-variable model;
A: The all (i.e., full) risk-variable model.

d. The number of risk variables in the model. This number sometimes changed across arrest transitions. A risk variable was included in the model at a particular arrest transition based on its distributional features. First, and most obviously, risk variables were included only when they could produce reliable estimates. For this reason, for example, the birth cohort subject's race during the juvenile period was included only at the first arrest transition; too few whites appeared at the later arrest transitions. Second, and less obviously, some variables were included in the model because of the way those variables were technically defined. For instance, the risk variable indicating the presence of an arrest for a prior UCR index crime was included only at the first arrest transition because the variable could take on different values only at this arrest transition. (The birth cohort subject might or might not have been previously arrested for a UCR property index crime.) At all subsequent transitions, the birth cohort subject must have had a prior arrest for a UCR index crime, at a minimum, the first arrest for a serious violent crime, which placed the subject in the study sample.

e. This value is minus two times the model's loglikelihood statistic, a statistic measuring how well the model matched the observed rearrest-time data. We have calculated minus two times this statistic because this value can be used to judge the comparative merits of nested risk-variable models, that is, of models whose coefficients are related in the form of superset to subset. One can compare, then, the "0" model to the "L," "L+R," and "A" models, the "L" model to the "L+R" and "A" models, and the "L+R" to the "A" model. These comparisons are transitive: if the "L" model is significantly superior to the "0" model, the "L+R" and the "A" models are also statistically superior because they include the "L" model.

When statistically comparing two risk-variable models, we performed the following steps: (1) calculated the difference between -2 times the loglikelihood of each model, which we have presented in the table, (2) calculated the degrees of freedom of the statistical test, which is the difference between the numbers of risk variables in the two models, (3) turned to a chi-square table, found the tabular cell entry which was at the intersection of the degrees-of-freedom and selected significance-level, and determined whether the value calculated in step 1 was greater than the tabular cell entry, and (4) reported that the comparison was statistically significant if the value calculated in step 1 was greater than the cell entry or, conversely, that the comparison was statistically nonsignificant if the opposite was true. An asterisk ("*") appearing after a loglikelihood value indicates that the broader risk-variable model (i.e., the superset) significantly improved (p. val. < .05) the explanatory capacity relative to the narrower risk-variable model (i.e., the subset) immediately adjacent to the left.

In some instances, the risk-variable model failed to converge. In such cases, it was impermissible to draw formal inferences about the magnitudes and signs of the model's coefficients. However, one can still loosely use the loglikelihood statistic of a nonconvergent model to compare the explanatory capacity of that model to other, related risk-variable models. We did not, however, formally compare a risk-variable model which failed to converge to any other model.

NC: The model did not converge.

f. The distribution's shape parameter defines the curvature of the hazard function. For some distributions, the shape parameter invariably produces a specific type of overall curvature, although the shape parameter's magnitude effects the details of that curvature. (For instance, the lognormal distribution is always single peaked, but the shape parameter governs the rate of incline to the peak and, in turn, the rate of decline thereafter.) For

Table 3.17

other distributions, the shape parameter produces generically different curvatures depending upon the magnitude of the parameter. The overall curvatures of the hazard functions of the Weibull and loglogistic distributions depend upon the magnitudes their shape parameters as follows:

Distribution	Value of Shape Parameter
Weibull (extreme value)	= 1, constant hazard function (i.e., the Weibull reduces to the exponential form),
	> 1, decreasing hazard function,
	< 1, increasing hazard function.
loglogistic	> 1, decreasing hazard function,
	$\leq$ 1, single-peaked hazard function.

An asterisk ("*") after the shape parameter indicates that the parameter was significantly different (p. val. < .05) from 1, in the indicated direction.

g. The "bullet" (".") before the variable name indicates the general name of the risk variable; the dash before the variable name indicates a specific level of the general risk variable.

h. [REF]: The omitted reference category of the categorical variable. The effect of a specific category of a categorical variable (e.g., the presence of a firearm) on the timing of rearrest is obtained by comparing the coefficient calculated for that category to the reference category (e.g., the absence of a weapon).

i. Only those coefficients are presented which were significant at p. val. < .05.

j. NE: The model was not estimated because there were too few cases to produce reliable results.

k. NS: Not significant at p. val. < .05.

l. NA: Not applicable. The risk variable or risk-variable model did not apply at the arrest transition.

Table 3.18

1958 Birth Cohort: Adult Arrests for Violent Crimes--
Failure-Time Regression-Model Loglikelihoods, Shape and
Scale Parameters, and Significant Risk Variables by
Arrest Transition, Race, Parametric Distribution,
and Risk Variable Model
(Construction Sample)

Panel A: 1st Arrest Transition--Total, Blacks, Whites
Panel B: 2nd Arrest Transition--Total, Blacks, Whites
Panel C: 3rd Arrest Transition--Blacks
Panel D: 4th Arrest Transition--Blacks
Panel E: 5th Arrest Transition--Blacks

Panel A.1: 1st Arrest Transition--Total (N = 911)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21	0	4	5	21
-2 Loglikelihood[e]	NA	4,182	4,162*	4,110*	6,062	6,038*	6,014*	5,989	5,890	5,868*	5,848*	5,802*	5,886	5,866*	5,846*	5,794*	5,894	5,876*	5,856*	5,802*
Shape[f]	NA	--	--	--	1.0	1.0	1.0	1.0	1.8*	1.8*	1.8*	1.8*	1.6*	1.6*	1.6*	1.6*	3.0	3.0	3.0	2.9
Scale[f]	NA	--	--	--	8.3	8.6	9.0	7.2	9.0	9.4	10.0	3.4	8.5	8.8	9.5	-2.8	8.7	9.1	9.8	-2.3
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	.4	.2	NS	NA	-.4	-.2	NS	NA	-.7	- .4	NS	NA	-.7	- .5	NS	NA	-.7	- .5	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Other Weapon	NA	-.4	-.4	- .4	NA	.4	.2	.4	NA	.7	.7	.7	NA	.8	.7	.7	NA	.8	.8	.7
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	.7	.7	NA	NA	-.7	-.7	NA	NA	-1.2	-1.3	NA	NA	-1.2	-1.3	NA	NA	-1.3	-1.3
.Age at Arrest for Present Violent Crime	NA	NA	NA	>-.1	NA	NA	NA	NA	NA	NA	NA	< .1	NA	NA	NS	< .1	NA	NA	NA	< .1

Table 3.18--Panel A.1 (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Other Weapon	NA	NA	NA	.7	NA	NA	NA	-1.0	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Socioeconomic Status < 15th Percentile	NA	NA	NA	NS	NA	NA	NA	- .2	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS

Table 3.18 (cont.)

Panel A.2: 1st Arrest Transition--Blacks (N = 693)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	20	0	4	NA	20	0	4	NA	20	0	4	NA	20	0	4	NA	20
-2 Loglikelihood[e]	NA	3,402	NA	3,364*	5,070	5,062	NA	5,036*	4,928	4,920	NA	4,882*	4,926	4,918	NA	4,876*	4,936	4,928	NA	4,888*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.8*	1.8*	NA	1.8*	1.6*	1.6*	NA	1.6*	2.9	2.9	NA	2.9
Scale[f]	NA	--	NA	--	8.2	8.0	NA	7.6	8.7	8.4	NA	3.6	8.1	7.9	NA	-2.3	8.3	8.0	NA	-2.5
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	-.4	NA	-.3	NA	.4	NA	.4	NA	NS	NA	.6	NA	NS	NA	.7	NA	.8	NA	.7
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Age at Arrest for Present Violent Crime	NA	NA	NA	>-.1	NA	NA	NA	< .1	NA	NA	NA	NS	NA	NA	NA	<.1	NA	NA	NA	<.1
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Other Weapon	NA	NA	NA	NS	NA	NA	NA	-1.2	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Socioeconomic Status ≤ 15th Percentile	NA	NA	NA	NS	NA	NA	NA	- .2	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS

Table 3.18 (cont.)

Panel A.3: 1st Arrest Transition--Whites (N = 218)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	NC	0	4	NA	NC	0	4	NA	NC	0	4	NA	NC	0	4	NA	NC
-2 Loglikelihood[e]	NA	480	NA	NC	956	942*	NA	NC	934	920*	NA	NC	932	920*	NA	NC	932	920*	NA	NC
Shape[f]	NA	--	NA	NC	1.0	1.0	NA	NC	1.8*	1.8*	NA	NC	1.7*	1.6*	NA	NC	3.3	3.2	NA	NC
Scale[f]	NA	--	NA	NC	9.0	9.7	NA	NC	10.1	11.3	NA	NC	9.7	10.8	NA	NC	10.1	11.3	NA	NC
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	.9	NA	NC	NA	-.9	NA	NC	NA	-1.6	NA	NC	NA	-1.6	NA	NC	NA	-1.6	NA	NC
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC
.Weapon Used																				
-Firearm	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC
-Other Weapon	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC	NA	NS	NA	NC
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.18 (cont.)

Panel B.1: 2nd Arrest Transition--Total (N = 325)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	25	0	4	5	25	0	4	5	25	0	4	5	25	0	4	5	25
-2 Loglikelihood[e]	NA	1,662	1,662	1,624*	2,736	2,718*	2,718	2,680*	2,672	2,658*	2,658	2,624*	2,670	2,658*	2,658	2,622*	2,674	2,660*	2,660	2,624*
Shape[f]	NA	--	--	--	1.0	1.0	1.0	1.0	1.7*	1.6*	1.6*	1.6*	1.4*	1.4*	1.4*	1.3*	2.6	2.5	2.5	2.4
Scale[f]	NA	--	--	--	7.7	7.2	7.3	9.5	8.0	7.3	7.4	8.7	7.4	6.8	6.7	8.4	7.5	6.8	6.5	7.4
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NS	NS	NA	NS	NS	-.2	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	-.3	-.3	-.3	NA	.3	.3	.4	NA	.5	.5	.5	NA	.5	.5	.5	NA	.5	.5	.5
.Weapon Used																				
-Firearm	NA	NS	NS	NS	NA	-.4	-.4	NS	NA	NS	NS	-.8	NA	NS	NS	-.9	NA	NS	NS	-1.0
-Other Weapon	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NS	NS	NA	NA	NS	NS	NA	NA	NS	NS	NA	NA	NS	NS	NA	NA	NS	NS
.Age at Arrest for Present Violent Crime	NA	NA	NA	>-.1	NA	NA	NA	NS	NA	NA	NA	<.1	NA	NA	NA	<.1	NA	NA	NA	< .1
.Socioeconomic Status < 15th Percentile	NA	NA	NA	.4	NA	NA	NA	-.6	NA	NA	NA	-.8	NA	NA	NA	-.7	NA	NA	NA	- .7

Table 3.18 (cont.)

Panel B.2: 2nd Arrest Transition--Blacks (N = 277)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model: Proportional Hazards[a] O[c]	L[c]	L+R[c]	A[c]	Exponential[b] O	L	L+R	A	Weibull[b] O	L	L+R	A	Loglogistic O	L	L+R	A	Lognormal O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	24	0	4	NA	24	0	4	NA	24	0	4	NA	24	0	4	NA	24
-2 Loglikelihood[e]	NA	1,402	NA	1,372*	2,382	2,370*	NA	2,338*	2,346	2,336*	NA	2,308*	2,342	2,336	NA	2,304*	2,348	2,338*	NA	2,308*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.5*	1.5*	NA	1.5*	1.3*	1.2*	NA	1.2*	2.3	2.3	NA	2.2
Scale[f]	NA	--	NA	--	7.7	7.4	NA	11.0	7.9	7.5	NA	10.6	7.3	7.1	NA	10.1	7.4	7.1	NA	9.8
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	-.3	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	.2	NA	.2	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Any Priors																				
-Yes	NA	NA	NA	NS	NA	NA	NA	-1.6	NA	NA	NA	-2.3	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-No [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Mean Seriousness																				
-Known Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	.4	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
.Incarcerated for a Prior UCR Index Crime	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	.9	NA	NA	NA	.9
.Socioeconomic Status ≤ 15th Percentile	NA	NA	NA	.5	NA	NA	NA	-.6	NA	NA	NA	-.7	NA	NA	NA	-.6	NA	NA	NA	-.7

Table 3.18 (cont.)

Panel B.3: 2nd Arrest Transition--Whites (N = 48)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE
-2 Loglikelihood[e]	NA	134	NA	NE	352	334	NA	NE	318	308*	NA	NE	316	306*	NA	NE	316	306*	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	2.7*	2.4*	NA	NE	2.3*	1.9*	NA	NE	4.0	3.3	NA	NE
Scale[f]	NA	--	NA	NE	7.8	5.5	NA	NE	8.8	5.1	NA	NE	7.9	4.0	NA	NE	8.0	4.0	NA	NE
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	-.9	NA	NE	NA	1.3	NA	NE	NA	2.2	NA	NE	NA	2.3	NA	NE	NA	2.3	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	-1.5	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	-1.0	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.18 (cont.)

Panel C: 3rd Arrest Transition--Blacks (N = 137)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	24	0	4	NA	24	0	4	NA	24	0	4	NA	24	0	4	NA	24
-2 Loglikelihood[e]	NA	612	NA	594	1,172	1,172	NA	1,154	1,160	1,160	NA	1,142	1,160	1,158	NA	1,140	1,164	1,162	NA	1,142
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.4*	1.4*	NA	1.4*	1.2*	1.2*	NA	1.1	2.2	2.2	NA	2.0
Scale[f]	NA	--	NA	--	7.5	7.4	NA	9.1	7.7	7.6	NA	8.8	7.1	7.0	NA	5.3	7.2	7.1	NA	6.2
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
First Prior UCR Index Crime																				
.Type																				
-Robbery	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	1.1	NA	NA	NA	NS
-Assault	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Property [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.18 (cont.)

Panel D: 4th Arrest Transition--Blacks (N = 69)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	24	0	4	NA	24	0	4	NA	24	0	4	NA	24	0	4	NA	24
-2 Loglikelihood[e]	NA	254	NA	218*	560	556	NA	536	542	538	NA	502	542	538	NA	500*	542	538	NA	502*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.8*	1.8*	NA	1.5*	1.5*	1.5*	NA	1.1	2.8	2.7	NA	2.1
Scale[f]	NA	--	NA	--	7.2	6.8	NA	6.9	7.6	6.8	NA	5.8	7.0	6.4*	NA	1.4	7.0	6.6	NA	1.4
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	1.8	NA	NS	NA	-1.0	NA	NS	NA	-2.3	NA	NS	NA	-2.3	NA	NS	NA	-2.3
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Age at First Arrest	NA	NA	NA	- .1	NA	NA	NA	.1	NA	NA	NA	.2	NA	NA	NA	.2	NA	NA	NA	.2
.Incarcerated for a Prior UCR Index Crime	NA	NA	NA	-1.6	NA	NA	NA	.8	NA	NA	NA	2.2	NA	NA	NA	2.8	NA	NA	NA	2.4
First Prior UCR Index Crime																				
.Age	NA	NA	NA	.1	NA	NA	NA	- .1	NA	NA	NA	- .2	NA	NA	NA	- .2	NA	NA	NA	- .2

Table 3.18--Panel D (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
Most Recent Prior UCR Index Crime																				
.Type																				
-Robbery	NA	NA	NA	1.3	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-2.4	NA	NA	NA	-2.3
-Assault	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-2.9	NA	NA	NA	-3.1
-Property [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	1.5	NA	NA	NA	1.6
-Other Weapon	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.18 (cont.)

Panel E: 5th Arrest Transition--Blacks (N = 34)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE
-2 Loglikelihood[e]	NA	104	NA	NE	290	286	NA	NE	274	272	NA	NE	274	270	NA	NE	274	270	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	2.0*	2.0*	NA	NE	1.7*	1.6*	NA	NE	3.1	2.8	NA	NE
Scale[f]	NA	--	NA	NE	7.0	6.1	NA	NE	7.4	6.3	NA	NE	6.6	6.2	NA	NE	6.7	6.3	NA	NE
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	-.9	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

a. At each arrest transition, a proportional hazards (i.e., Cox) regression model was estimated in addition to the four parametric failure time regression models. This model permits one to estimate the effects of risk variables on the time-specific rearrest risk (i.e., the hazard rate) but, because it makes no assumption about the shape of the underlying parametric distribution which generated these risks, does not permit one to estimate distributional parameters. (Thus, it is a semiparametric model--parametric only in the coefficients.) The proportional hazards model assumes that the hazard functions of different levels, or strata, of a risk variable (e.g., low SES and high SES) are proportional to one another across time. For instance, low SES subjects might have a hazard rate that is twice as high as high SES subjects regardless of the point in time at which one contrasts the groups. While the proportional hazards model might have limited use with respect to js policy making which aims to devise interventions which are linked to time, because the model does not assign a specific time-varying shape to the hazard function, it does have an important feature: it is robust. Risk-variable coefficients are reliably estimated by the proportional hazards model across a variety of parametric distributions (e.g., exponential, Weibull, loglogistic). Because it is robust, the proportional hazards model can be used as a baseline model against which to compare the coefficient estimates produced by the other parametric models. We use the proportional hazards model, therefore, to check the consistency and, in turn, plausibility of our results.

As the name hints, the risk variables in the proportional hazards model influence the <u>hazard function</u>, which is the rearrest function in this study. A positive coefficient indicates that the presence of a risk characteristic increases the rearrest function; a negative coefficient indicates the reverse. To compare the effect of a coefficient estimated by the proportional hazards model to the effect of a coefficient estimated by one of the other parametric models, one simply flips the sign of the proportional hazards coefficient. This is done because the parametric models produce coefficient estimates based on the relationship of the risk variable to the timing of rearrest, not to the rearrest (i.e., hazard) rate. When a coefficient in the proportional

hazards model has a positive sign, the risk variable increases the rearrest risk, which corresponds to a negative sign in the parametric models, indicating a more rapid time until rearrest. Put somewhat differently, a higher hazard rate implies a shorter rearrest time, indicated by the opposite signs of the same coefficient in the proportional hazards and the parametric models.

To estimate a proportional hazards model, at least one risk variable must be included. For this reason, the unconditional risk-variable model (i.e., no risk variables included), designated by the tabular column "0," does not apply and is, therefore, not filled in. Also, because the proportional hazards model is semiparametric in the way described above, the tabular rows designating the "shape" and "scale" parameters do not apply and, likewise, are not filled in. See notes c and f for further discussion of the different risk-variable models which were estimated and of some special features of the shape and scale parameters.

b. We estimated the extreme value parameterization of this model. For details of the computational procedure, see D. Steinberg and P. Colla, SURVIVAL: A Supplementary Module for SYSTAT (Evanston, IL: SYSTAT, Inc, 1988). The extreme value parameterization yields coefficient estimates which are identical in magnitude and sign to those estimated by the nonextreme value parameterizations. Only the shape and scale parameters differ across parameterizations, and these are easily converted to one another. Technical estimation issues mainly influenced the decision to use this parameterization.

c. 0: The unconditional model (i.e., no risk variables included);
L: The legally-permissible risk-variable model;
L+R: The legally-permissible-plus-race risk-variable model;
A: The all (i.e., full) risk-variable model.

d. The number of risk variables in the model. This number sometimes changed across arrest transitions. A risk variable was included in the model at a particular arrest transition based on its distributional features. First, and most obviously, risk variables were included only when they could produce reliable estimates. For this reason, for example, the birth cohort subject's race during the juvenile period was included only at the first arrest transition; too few whites appeared at the later arrest transitions. Second, and less obviously, some variables were included in the model because of the way those variables were technically defined. For instance, the risk variable indicating the presence of an arrest for a prior UCR index crime was included only at the first arrest transition because the variable could take on different values only at this arrest transition. (The birth cohort subject might or might not have been previously arrested for a UCR property index crime.) At all subsequent transitions, the birth cohort subject must have had a prior arrest for a UCR index crime, at a minimum, the first arrest for a serious violent crime, which placed the subject in the study sample.

e. This value is minus two times the model's loglikelihood statistic, a statistic measuring how well the model matched the observed rearrest-time data. We have calculated minus two times this statistic because this value can be used to judge the comparative merits of nested risk-variable models, that is, of models whose coefficients are related in the form of superset to subset. One can compare, then, the "0" model to the "L," "L+R," and "A" models, the "L" model to the "L+R" and "A" models, and the "L+R" to the "A" model. These comparisons are transitive: if the "L" model is significantly superior to the "0" model, the "L+R" and the "A" models are also statistically superior because they include the "L" model.

When statistically comparing two risk-variable models, we performed the following steps: (1) calculated the difference between -2 times the loglikelihood of each model, which we have presented in the table, (2) calculated the degrees of freedom of the statistical test, which is the difference between the numbers of risk variables in the two models, (3) turned to a chi-square table, found the tabular cell entry which was at the intersection of the degrees-of-freedom and selected significance-level, and determined whether the value calculated in step 1 was greater than the tabular cell entry, and (4) reported that the comparison was statistically significant if the value calculated in step 1 was greater than the cell entry or, conversely, that the comparison was statistically nonsignificant if the opposite was true. An asterisk ("*") appearing after a loglikelihood value indicates that the broader risk-variable model (i.e., the superset) significantly improved (p. val. < .05) the explanatory capacity relative to the narrower risk-variable model (i.e., the subset) immediately adjacent to the left.

In some instances, the risk-variable model failed to converge. In such cases, it was impermissible to draw formal inferences about the magnitudes and signs of the model's coefficients. However, one can still loosely use the loglikelihood statistic of a nonconvergent model to compare the explanatory capacity of that model to other, related risk-variable models. We did not, however, formally compare a risk-variable model which failed to converge to any other model.

NC: The model did not converge.

f. The distribution's shape parameter defines the curvature of the hazard function. For some distributions, the shape parameter invariably produces a specific type of overall curvature, although the shape parameter's magnitude effects the details of that curvature. (For instance, the lognormal distribution is always single peaked, but the shape parameter governs the rate of incline to the peak and, in turn, the rate of decline thereafter.) For

other distributions, the shape parameter produces generically different curvatures depending upon the magnitude of the parameter. The overall curvatures of the hazard functions of the Weibull and loglogistic distributions depend upon the magnitudes their shape parameters as follows:

Distribution	Value of Shape Parameter
Weibull (extreme value)	= 1, constant hazard function (i.e., the Weibull reduces to the exponential form), > 1, decreasing hazard function, < 1, increasing hazard function.
loglogistic	> 1, decreasing hazard function, ≤ 1, single-peaked hazard function.

An asterisk ("*") after the shape parameter indicates that the parameter was significantly different (p. val. < .05) from 1, in the indicated direction.

g. The "bullet" (".") before the variable name indicates the general name of the risk variable; the dash before the variable name indicates a specific level of the general risk variable.

h. [REF]: The omitted reference category of the categorical variable. The effect of a specific category of a categorical variable (e.g., the presence of a firearm) on the timing of rearrest is obtained by comparing the coefficient calculated for that category to the reference category (e.g., the absence of a weapon).

i. Only those coefficients are presented which were significant at p. val. < .05.

j. NE: The model was not estimated because there were too few cases to produce reliable results.

k. NS: Not significant at p. val < .05.

l. NA: Not applicable. The risk variable or risk-variable model did not apply at the arrest transition.

Table 3.19

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Selected Observed Rearrest-Time (in Months) Percentiles
by Age Status and Race; Exposed for the Juvenile and
Young Adult Years--Ages 10-26
(Construction Sample)

Age Status and Race	1st					2nd					3rd					4th					5th				
	(N)[a]	10	25	50	90	(N)	10	25	50	90	(N)	10	25	50	90	(N)	10	25	50	90	(N)	10	25	50	90
Juveniles																									
Total	(759)	4	15	92	NA[b]	(262)	2	7	25	NA	(124)	1	3	19	NA	(62)	1	4	11	NA	(37)	<.5	1	7	NA
Blacks	(644)	3	13	69	NA	(245)	2	7	24	NA	(117)	1	3	18	NA	(59)	1	4	11	NA	(36)	<.5	1	7	NA
Whites	(115)	12	111	NA	NA	(17)	--[c]	--	--	--		--	--	--	--		--	--	--	--		--	--	--	--

a. The number of birth cohort subjects at risk of rearrest.

b. The cell entry is not applicable because the percentile was not reached.

c. There were too few cases (N $\leq$ 30) to compute the rearrest time percentile.

Table 3.20

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles by
Type of Distribution and Arrest Transition;
Exposed for the Juvenile and Young Adult
Years--Ages 10-26
(Construction Sample)

	Arrest Transition																			
	1st				2nd				3rd				4th				5th			
Distribution	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90
Observed	4	15	92	NA[a]	2	7	25	NA	1	3	19	NA	1	4	11	NA	<.5	1	7	NA
Exponential	16	43	103	341	7	20	48	160	7	18	44	148	5	12	30	100	4	10	25	82
Split Exponential	7	20	81	NA	3	10	27	NA	3	7	20	NA	2	5	14	NA	1	3	8	NA
Loglogistic	4	18	96	2,659	2	7	28	485	1	4	19	541	1	3	13	196	<.5	2	7	140
Split Loglogistic	4	16	95	NA	2	6	24	NA	1	4	15	NA	1	3	11	NA	1	2	5	NA
Lognormal	4	17	99	2,717	2	7	29	473	1	4	20	509	1	3	14	205	<.5	2	8	153
Split Lognormal	4	15	98	NA	2	6	24	NA	1	3	15	NA	1	3	11	NA	1	2	5	NA
Weibull	3	20	103	945	1	7	34	295	<.5	4	26	327	<.5	3	17	163	<.5	2	10	143
Split Weibull	4	17	90	NA	2	7	25	NA	1	4	17	NA	1	4	12	NA	<.5	2	6	NA
Gompertz	6	19	81	NA	3	8	24	NA	2	6	17	NA	2	4	12	NA	1	2	7	NA
Mixed Exponential	5	16	101	845	2	7	23	372	2	5	16	494	2	5	12	453	--[b]	--	--	--

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

b. The rearrest percentile was not reported for this split population distribution because the results were identical to those for the related unitary-population distribution. Identical results can occur for the split forms of the exponential loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that these findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.21

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Selected Rearrest-Time (in Months) Percentiles by Type
of Distribution, Race, and Arrest Transition;
Exposed for the Juvenile and Young Adult
Years--Ages 10-26
(Construction Sample)

	Arrest Transition																							
	Blacks																				Whites			
	1st				2nd				3rd				4th				5th				1st			
Distribution	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90	10	25	50	90
Observed	3	13	69	NA[a]	2	7	24	NA	1	3	18	NA	1	4	11	NA	<.5	1	7	NA	12	111	NA	NA
Exponential	14	37	89	294	7	20	48	158	7	20	47	156	5	12	29	98	4	11	25	84	42	114	275	913
Split Exponential	6	18	60	NA	3	10	27	NA	3	7	21	NA	2	5	14	NA	1	3	8	NA	17	86	NA	NA
Loglogistic	3	15	73	1,778	2	7	27	461	1	4	20	636	1	3	13	193	<.5	2	7	155	15	105	713	33,000
Split Loglogistic	4	13	68	NA	2	6	23	NA	1	3	16	NA	1	3	11	NA	1	2	6	NA	13	95	NA	NA
Lognormal	3	14	75	1,779	2	7	28	449	1	4	22	594	1	3	13	199	<.5	2	8	167	15	104	908	56,000
Split Lognormal	3	13	70	NA	2	6	24	NA	1	3	16	NA	1	3	11	NA	1	2	5	NA	12	98	NA	NA
Weibull	3	17	81	715	1	7	34	285	<.5	4	27	372	<.5	3	17	157	<.5	2	11	151	16	109	583	5,757
Split Weibull	4	14	65	NA	2	7	25	NA	1	4	18	NA	1	4	12	NA	<.5	2	7	NA	14	93	NA	NA
Gompertz	5	16	59	NA	3	8	24	NA	2	6	17	NA	2	4	12	NA	1	2	7	NA	16	86	NA	NA
Mixed Exponential	4	14	71	670	2	7	23	336	2	5	16	504	2	5	13	434	--[b]	--	--	--	14	98	699	3,104

a. The cell entry is not applicable because the designated distribution indicated that this percentile was not reached.

b. The rearrest percentile was not reported for this split population distribution because the results were identical to those for the related unitary-population distribution. Identical results can occur for the split forms of the exponential loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that these findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.22

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Observed Monthly Hazard Rates by Race and Arrest
Transition; Exposed for the Juvenile and
Young Adult Years--Ages 10-26
(Construction Sample)

	Total					Blacks					Whites
Month	1st (N = 759)[a]	2nd (N = 262)	3rd (N = 124)	4th (N = 62)	5th (N = 37)	1st (N = 644)	2nd (N = 245)	3rd (N = 117)	4th (N = 59)	5th (N = 36)	1st (N = 115)
1	.046	.071	.157	.157	.277	.051	.068	.157	.165	.286	.018
2	.017	.046	.068	.019	.036	.020	.049	.073	.020	.038	.000
3	.020	.044	.031	.019	.118	.022	.042	.033	.021	.080	.009
4	.033	.023	.076	.103	.133	.038	.024	.069	.087	.133	.009
5	.018	.047	.058	.022	.049	.021	.045	.061	.023	.049	.000
6	.015	.019	.012	.143	.000	.011	.021	.013	.150	.000	.037
7	.012	.030	.012	.026	.105	.011	.032	.013	.027	.105	.019
8	.014	.052	.050	.027	.057	.015	.056	.040	.028	.057	.010
9	.024	.066	.053	.085	.061	.029	.071	.056	.059	.061	.000
10	.008	.035	.000	.030	.133	.010	.038	.000	.031	.133	.000
11	.016	.024	.041	.095	.000	.020	.019	.044	.098	.000	.000
12	.010	.006	.014	.069	.000	.012	.007	.015	.071	.000	.000
13	.020	.025	.029	.000	.154	.020	.026	.031	.000	.154	.019
14	.017	.044	.015	.000	.087	.019	.048	.016	.000	.087	.010
15	.019	.007	.030	.074	.000	.023	.007	.032	.077	.000	.000
16	.002	.026	.000	.039	.000	.002	.022	.000	.041	.000	.000
17	.011	.027	.000	.041	.000	.013	.029	.000	.043	.000	.000
18	.009	.007	.031	.043	.000	.011	.007	.033	.044	.000	.000
19	.007	.014	.032	.000	.000	.009	.015	.034	.000	.000	.000
20	.009	.014	.050	.000	.000	.011	.015	.036	.000	.000	.000
21	.009	.022	.017	.044	.000	.011	.016	.018	.047	.000	.000
22	.011	.015	.000	.000	.000	.014	.016	.000	.000	.000	.000
23	.006	.015	.018	.000	.095	.007	.016	.019	.000	.095	.000
24	.009	.008	.036	.000	.000	.009	.008	.038	.000	.000	.010
25	.006	.015	.000	.047	.000	.007	.017	.000	.049	.000	.000
26	.015	.000	.000	.000	.000	.019	.000	.000	.000	.000	.000
27	.006	.008	.019	.000	.000	.005	.008	.020	.000	.000	.010
28	.006	.016	.000	.000	.000	.007	.008	.000	.000	.000	.000
29	.006	.000	.019	.000	.000	.007	.000	.020	.000	.000	.000
30	.012	.048	.000	.000	.105	.012	.052	.000	.000	.105	.010
31	.010	.008	.000	.049	.000	.010	.009	.000	.051	.000	.010
32	.008	.008	.019	.051	.000	.010	.009	.021	.054	.000	.000
33	.012	.017	.020	.000	.118	.010	.009	.021	.000	.118	.021
34	.010	.009	.020	.000	.000	.013	.000	.000	.000	.000	.000
35	.004	.017	.000	.054	.000	.005	.019	.000	.057	.000	.000
36	.006	.009	.021	.000	.000	.008	.009	.022	.000	.000	.000
37	.006	.009	.021	.000	.000	.008	.009	.022	.000	.000	.000
38	.002	.000	.043	.000	.000	.003	.000	.045	.000	.000	.000
39	.006	.009	.022	.000	.000	.005	.010	.024	.000	.000	.011
40	.004	.018	.023	.057	.000	.005	.019	.024	.061	.000	.000
41	.000	.009	.024	.061	.000	.000	.010	.000	.065	.000	.000

Table 3.22 (cont.)

	Total					Blacks					Whites
Month	1st (N = 759)	2nd (N = 262)	3rd (N = 124)	4th (N = 62)	5th (N = 37)	1st (N = 644)	2nd (N = 245)	3rd (N = 117)	4th (N = 59)	5th (N = 36)	1st (N = 115)
42	.006	.028	.049	.000	.133	.008	.030	.025	.000	.133	.000
43	.002	.019	.000	.000	.000	.003	.021	.000	.000	.000	.000
44	.000	.010	.000	.000	.000	.000	.000	.000	.000	.000	.000
45	.002	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000
46	.009	.010	.000	.065	.000	.005	.010	.000	.069	.000	.022
47	.002	.000	.025	.000	.000	.003	.000	.025	.000	.000	.000
48	.013	.020	.000	.000	.000	.017	.021	.000	.000	.000	.000
49	.004	.000	.000	.000	.000	.006	.000	.000	.000	.000	.000
50	.007	.020	.000	.069	.000	.008	.022	.000	.074	.000	.000
51	.004	.000	.000	.000	.000	.006	.000	.000	.000	.000	.000
52	.002	.010	.000	.000	.000	.003	.011	.000	.000	.000	.000
53	.009	.000	.000	.000	.000	.006	.000	.000	.000	.000	.022
54	.007	.032	.000	.074	.000	.009	.034	.000	.080	.000	.000
55	.007	.011	.000	.000	.000	.006	.012	.000	.000	.000	.011
56	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
57	.007	.000	.000	.000	.000	.009	.000	.000	.000	.000	.000
58	.005	.000	.000	.000	.000	.006	.000	.000	.000	.000	.000
59	.012	.000	.000	.000	.000	.015	.000	.000	.000	.000	.000
60	.005	.000	.000	.000	.000	.006	.000	.000	.000	.000	.000
61	.000	.011	.000	.000	.000	.000	.012	.000	.000	.000	.000
62	.002	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000
63	.002	.000	.026	.000	.000	.003	.000	.026	.000	.000	.000
64	.007	.000	.000	.000	.000	.006	.000	.000	.000	.000	.011
65	.005	.011	.000	.000	.000	.006	.012	.000	.000	.000	.000
66	.005	.000	.000	.000	.000	.006	.000	.000	.000	.000	.000
67	.002	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000
68	.000	.011	.000	.000	.000	.000	.012	.000	.000	.000	.000
69	.007	.000	.000	.000	.000	.009	.000	.000	.000	.000	.000
70	.005	.000	.000	.000	.000	.006	.000	.000	.000	.000	.000
71	.012	.000	.000	.000	.000	.013	.000	.000	.000	.000	.011
72	.002	.011	.000	.000	.000	.003	.012	.000	.000	.000	.000
73	.003	.011	.000	.000	.000	.003	.012	.000	.000	.000	.000
74	.000	.000	.027	.000	.000	.000	.000	.027	.000	.000	.000
75	.000	.012	.000	.000	.000	.000	.012	.000	.000	.000	.000
76	.003	.012	.000	.000	.000	.003	.013	.000	.000	.000	.000
77	.003	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000
78	.005	.012	.000	.000	.000	.006	.013	.000	.000	.000	.000
79	.008	.000	.000	.000	.000	.010	.000	.000	.000	.000	.000
80	.003	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000
81	.003	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000
82	.003	.000	.000	.000	.000	.003	.000	.000	.000	.000	.000
83	.005	.012	.000	.000	.000	.007	.013	.000	.000	.000	.000
84	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
85	.005	.000	.000	.000	.000	.007	.000	.000	.000	.000	.000
86	.005	.000	.000	.000	.000	.007	.000	.000	.000	.000	.000
87	.003	.000	.000	.080	.000	.003	.000	.000	.087	.000	.000
88	.000	.012	.027	.000	.000	.000	.013	.027	.000	.000	.000

Table 3.22 (cont.)

	Total					Blacks					Whites
Month	1st (N = 759)	2nd (N = 262)	3rd (N = 124)	4th (N = 62)	5th (N = 37)	1st (N = 644)	2nd (N = 245)	3rd (N = 117)	4th (N = 59)	5th (N = 36)	1st (N = 115)
89	.005	.000	.000	.000	.000	.007	.000	.000	.000	.000	.000
90	.003	.025	.000	.000	.000	.003	.027	.000	.000	.000	.000
91	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
92	.000	.025	.000	.000	.000	.000	.027	.000	.000	.000	.000
93	.005	.000	.000	.000	.000	.007	.000	.000	.000	.000	.000
94	.003	.026	.000	.000	.000	.003	.028	.000	.000	.000	.000
95	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
96	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

a. The number of birth cohort subjects at risk of rearrest.

Table 3.23

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Loglikelihood Statistic by Type of Parametric
Distribution, Race, and Arrest Transition;
Exposed for the Juvenile and Young
Adult Years--Ages 10-26
(Construction Sample)

		Arrest Transition									
		Total									
		1st (N = 759)[a] (R = 407)[b]		2nd (N = 262) (R = 190)		3rd (N = 124) (R = 91)		4th (N = 62) (R = 50)		5th (N = 37) (R = 30)	
Parametric Distribution	Number of Parameters[c]	LL[d]	% R[e]	LL	% R	LL	% R	LL	% R	LL	% R
Exponential	1	-2441	100	-996	100	-470	100	-238	100	-137	100
Split Exponential	2	-2315	55	-934	73	-423	73	-213	81	-110	81
Loglogistic	2	-2309	100	-923	100	-413	100	-213	100	-112	100
Split Loglogistic	3	-2297	67	-922	82	409	81	-211	85	-108	83
Lognormal	2	-2300	100	-925	100	-411	100	-213	100	-111	100
Split Lognormal	3	-2295	71	-921	82	-407	79	-210	84	-107	82
Weibull	2	-2321	100	-939	100	-420	100	-218	100	-116	100
Split Weibull	3	-2298	57	-925	74	-411	74	-210	81	-108	81
Gompertz	2	-2312	56	-930	74	-418	75	-212	82	-110	82
Mixed Exponential	3	-2305	100	-925	100	-415	100	-212	100	--[f]	--

Table 3.23 (cont.)

Parametric Distribution	Number of Parameters[c]	Arrest Transition: Blacks 1st (N = 644)[a] (R = 376)[b] LL[d]	% R[e]	Blacks 2nd (N = 245) (R = 179) LL	% R	Blacks 3rd (N = 117) (R = 84) LL	% R	Blacks 4th (N = 59) (R = 48) LL	% R	Blacks 5th (N = 36) (R = 29) LL	% R	Whites 1st (N = 115) (R = 31) LL	% R
Exponential	1	-2199	100	-936	100	-438	100	-228	100	-133	100	-216	100
Split Exponential	2	-2088	59	-880	73	-392	72	-204	81	-108	81	-206	28
Loglogistic	2	-2081	100	-873	100	-382	100	-204	100	-109	100	-207	100
Split Loglogistic	3	-2072	73	-868	82	-378	78	-203	86	-106	83	-205	35
Lognormal	2	-2074	100	-870	100	-380	100	-204	100	-109	100	-206	100
Split Lognormal	3	-2070	77	-866	82	-376	77	-202	85	-105	81	-205	38
Weibull	2	-2093	100	-883	100	-388	100	-209	100	-113	100	-207	100
Split Weibull	3	-2073	62	-871	75	-380	73	-202	82	-106	81	-205	29
Gompertz	2	-2086	61	-875	75	-388	74	-204	82	-107	82	-206	28
Mixed Exponential	3	-2080	100	-870	100	-383	100	-204	100	--	--	-205	100

a. The number of birth cohort subjects at risk of rearrest.

b. The number of rearrested birth cohort subjects.

c. The number of parameters characterizing the distribution.

d. The distribution loglikelihood statistic.

e. The percentage of birth cohort subjects estimated to be rearrested by the distribution.

f. The loglikelihood statistic and corresponding percentage rearrested were not reported for this split-population distribution because the results were identical to the related unitary-population distribution. Identical results can occur for the split forms of the exponential, loglogistic, lognormal, and Weibull distributions. With respect to the mixed exponential distribution, nonreported findings indicate that the findings were identical to those of either the unitary- or split-population forms of the exponential distribution.

Table 3.24

1958 Birth Cohort: Juvenile Arrests for Violent Crimes--
Failure-Time Regression-Model Loglikelihoods, Shape and
Scale Parameters, and Significant Risk Variables by
Arrest Transition, Race, Parametric Distribution,
and Risk Variable Model; Exposed for the Juvenile
and Young Adult Years--Ages 10-26
(Construction Sample)

Panel A: 1st Arrest Transition--Total, Blacks, Whites
Panel B: 2nd Arrest Transition--Blacks
Panel C: 3rd Arrest Transition--Blacks
Panel D: 4th Arrest Transition--Blacks
Panel E: 5th Arrest Transition--Blacks

Panel A.1: 1st Arrest Transition--Total (N = 759)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	22	0	4	5	22	0	4	5	22	0	4	5	22	0	4	5	22
-2 Loglikelihood[e]	NA	5,016	5,072*	5,016*	7,662	7,644*	7,602*	7,524*	7,382	7,370*	7,336*	7,278*	7,366	7,356*	7,332*	7,262*	7,372	7,364*	7,332*	7,280*
Shape[f]	NA	--	--	--	1.0	1.0	1.0	1.0	1.9*	1.9*	1.9*	1.8*	1.6*	1.6*	1.6*	1.5*	2.9	2.9	2.8	2.7
Scale[f]	na	--	--	--	8.4	9.2	9.9	9.9	8.8	10.1	11.3	10.6	8.0	9.3	10.6	9.9	8.1	9.3	10.6	10.2
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	.3	NS	NS	NA	-.4	- .2	- .2	NA	-.6	NS	NS	NA	-.6	NS	NS	NA	-.5	NS	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	.2	NS	NS	NA	-.2	- .2	- .2	NA	-.4	NS	NS	NA	-.4	NS	NS	NA	-.4	NS	NS
.Weapon Used																				
-Firearm	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Other Weapon	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	1.0	.9	NA	NA	-1.1	-1.0	NA	NA	-1.9	-1.7	NA	NA	-1.9	-1.8	NA	NA	-2.0	-1.8
.Prior Status Offense	NA	NA	NA	.5	NA	NA	NA	- .6	NA	NA	NA	- .9	NA	NA	NA	- .9	NA	NA	NA	-1.0

Table 3.24--Panel A.1 (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Any Priors																				
-Yes	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown	NA	NA	NA	NS	NA	NA	NA	-1.2	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-No [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
First Prior UCR Index Crime																				
.Seriousness (Log)	NA	NA	NA	.5	NA	NA	NA	- .7	NA	NA	NA	-1.0	NA	NA	NA	-1.1	NA	NA	NA	NS
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	- .6	NA	NA	NA	- .8	NA	NA	NA	NS	NA	NA	NA	NS
-Other Weapon	NA	NA	NA	.4	NA	NA	NA	- .4	NA	NA	NA	- .7	NA	NA	NA	- .7	NA	NA	NA	- .8
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.24 (cont.)

Panel A.2: 1st Arrest Transition--Blacks (N = 644)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	0[c]	L[c]	L+R[c]	A[c]	0	L	L+R	A	0	L	L+R	A	0	L	L+R	A	0	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21
-2 Loglikelihood[e]	NA	4,570	NA	4,524*	6,966	6,958	NA	6,894*	6,720	6,714	NA	6,666*	6,704	6,700	NA	6,652*	6,712	6,710	NA	6,666*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.9*	1.9*	NA	1.8*	1.5*	1.5*	NA	1.5*	2.8	2.8	NA	2.6
Scale[f]	NA	--	NA	--	8.3	8.8	NA	9.0	8.5	9.4	NA	9.1	7.7	8.5	NA	8.7	7.8	8.5	NA	8.7
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	-.2	NA	-.2	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	-.2	NA	-.2	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Prior Status Offense	NA	NA	NA	.5	NA	NA	NA	-.6	NA	NA	NA	-.9	NA	NA	NA	-.9	NA	NA	NA	-.9
Prior Arrests for UCR Index Crimes																				
.Mean Seriousness	NA	NA	NA	NS	NA	NA	NA	1.0	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
First Prior UCR Index Crime																				
.Seriousness (Log)	NA	NA	NA	.6	NA	NA	NA	-.8	NA	NA	NA	-1.2	NA	NA	NA	-1.4	NA	NA	NA	-1.5
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	-.5	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Other Weapon	NA	NA	NA	NS	NA	NA	NA	-.4	NA	NA	NA	-.6	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.24 (cont.)

Panel A.3: 1st Arrest Transition--Whites (N = 115)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21
-2 Loglikelihood[e]	NA	276	NA	246*	644	636	NA	590*	622	614	NA	582*	622	614	NA	586*	622	614	NA	592
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	2.0*	2.0*	NA	1.6*	1.9*	1.8*	NA	1.4	3.6	3.5	NA	2.9
Scale[f]	NA	--	NA	--	9.4	9.9	NA	10.1	10.6	11.6	NA	10.3	10.1	11.3	NA	9.7	10.5	11.6	NA	9.0
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	.9	NA	NS	NA	-.9	NA	NS	NA	-1.8	NA	NS	NA	-1.9	NA	NS	NA	-1.8	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Socioeconomic Status <15th Percentile	NA	NA	NA	1.5	NA	NA	NA	-1.8	NA	NA	NA	NS	NA	NA	NA	-2.5	NA	NA	NA	NS

Table 3.24 (cont.)

Panel B: 2nd Arrest Transition--Blacks (N = 245)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25
-2 Loglikelihood[e]	NA	1,782	NA	1,742*	3,094	3,090*	NA	3,010	2,976	2,972	NA	2,928*	2,960	2,956	NA	2,922*	2,964	2,960	NA	2,926*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.8*	1.8*	NA	1.7*	1.4*	1.3*	NA	1.2*	2.4	2.4	NA	2.2
Scale[f]	NA	--	NA	--	7.6	7.1	NA	1.6	7.6	6.6	NA	< .1	6.7	5.2	NA	- .4	6.7	5.2	NA	- .8
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	.2	NA	NS	NA	NS	NA	NS	NA	.5	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	- .2	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	.4	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Age at Arrest for Present Violent Crime	NA	NA	NA	>- .1	NA	NA	NA	< .1	NA	NA	NA	< .1	NA	NA	NA	<.1	NA	NA	NA	<.1
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Any Priors																				
-Yes	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown	NA	NA	NA	1.9	NA	NA	NA	-3.0	NA	NA	NA	-3.3	NA	NA	NA	NS	NA	NA	NA	NS
-No [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Mean Seriousness																				
-Known Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	1.0	NA	NA	NA	NS	NA	NA	NA	1.0	NA	NA	NA	NS

Table 3.24--Panel B (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
First Prior UCR Index Crime																				
.Type																				
-Robbery	NA	NA	NA	NS	NA	NA	NA	.7	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Assault	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS				
-Property [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Other Weapon	NA	NA	NA	.3	NA	NA	NA	- .4	NA	NA	NA	NS	NA	NA	NA	-.7	NA	NA	NA	-.6
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.24 (cont.)

Panel C: 3rd Arrest Transition--Blacks (N = 117)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25
-2 Loglikelihood[e]	NA	712	NA	698	1,450	1,446	NA	1,400*	1,334	1,330	NA	1,312	1,326	1,322	NA	1,308	1,324	1,320	NA	1,308
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	2.3*	2.3*	NA	2.1*	1.7*	1.7*	NA	1.6*	2.9	2.9	NA	2.7
Scale[f]	NA	--	NA	--	7.6	6.7	NA	3.1	7.5	6.2	NA	.2	6.4	5.1	NA	- .6	6.4	5.5	NA	.2
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	.3	NA	.4	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Age at First Arrest	NA	NA	NA	>-.1	NA	NA	NA	< .1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
Prior Arrests for UCR Index Crimes																				
.Mean Seriousness	NA	NA	NA	NS	NA	NA	NA	1.1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Any Priors																				
-Yes	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown	NA	NA	NA	NS	NA	NA	NA	3.5	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-No [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.24--Panel C (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
.Mean Seriousness																				
-Known Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	-1.3	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
First Prior UCR Index Crime																				
.Age	NA	NA	NA	<.1	NA	NA	NA	>- .1	NA	NA	NA	>-.1	NA	NA	NA	>-.1	NA	NA	NA	>-.1

Table 3.24 (cont.)

Panel D: 4th Arrest Transition--Blacks (N = 59)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE
-2 Loglikelihood[e]	NA	330	NA	NE	784	776	NA	NE	738	732	NA	NE	732	726	NA	NE	734	726	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	2.0*	1.9*	NA	NE	1.4*	1.3*	NA	NE	2.5	2.3	NA	NE
Scale[f]	NA	--	NA	NE	7.2	8.0	NA	NE	6.9	8.0	NA	NE	5.9	7.2	NA	NE	5.9	7.1	NA	NE
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NE	NA	-.7	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	-.5	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	< .1	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.24 (cont.)

Panel E: 5th Arrest Transition--Blacks (N = 36)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE
-2 Loglikelihood[e]	NA	170	NA	NE	464	448*	NA	NE	420	416	NA	NE	414	410	NA	NE	414	410	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	2.3*	2.1*	NA	NE	1.5*	1.4	NA	NE	2.5	2.3	NA	NE
Scale[f]	NA	--	NA	NE	7.0	12.4	NA	NE	6.6	12.6	NA	NE	5.4	9.7	NA	NE	5.5	10.4	NA	NE
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[e]																				
-Robbery[g]	NA	NS	NA	NE	NA	-2.4	NA	NE	NA	-2.5	NA	NE	NA	-2.0	NA	NE	NA	-2.2	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NE	NA	-1.4	NA	NE	NA	-1.6	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	.4	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	-1.5	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

a. At each arrest transition, a proportional hazards (i.e., Cox) regression model was estimated in addition to the four parametric failure time regression models. This model permits one to estimate the effects of risk variables on the time-specific rearrest risk (i.e., the hazard rate) but, because it makes no assumption about the shape of the underlying parametric distribution which generated these risks, does not permit one to estimate distributional parameters. (Thus, it is a semiparametric model--parametric only in the coefficients.) The proportional hazards model assumes that the hazard functions of different levels, or strata, of a risk variable (e.g., low SES and high SES) are proportional to one another across time. For instance, low SES subjects might have a hazard rate that is twice as high as high SES subjects regardless of the point in time at which one contrasts the groups. While the proportional hazards model might have limited use with respect to js policy making which aims to devise interventions which are linked to time, because the model does not assign a specific time-varying shape to the hazard function, it does have an important feature: it is robust. Risk-variable coefficients are reliably estimated by the proportional hazards model across a variety of parametric distributions (e.g., exponential, Weibull, loglogistic). Because it is robust, the proportional hazards model can be used as a baseline model against which to compare the coefficient estimates produced by the other parametric models. We use the proportional hazards model, therefore, to check the consistency and, in turn, plausibility of our results.

As the name hints, the risk variables in the proportional hazards model influence the <u>hazard function</u>, which is the rearrest function in this study. A positive coefficient indicates that the presence of a risk characteristic increases the rearrest function; a negative coefficient indicates the reverse. To compare the effect of a coefficient estimated by the proportional hazards model to the effect of a coefficient estimated by one of the other parametric models, one simply flips the sign of the proportional hazards coefficient. This is done because the parametric models produce coefficient estimates based on the relationship of the risk variable to the timing of rearrest, not to the rearrest (i.e., hazard) rate. When a coefficient in the proportional

hazards model has a positive sign, the risk variable increases the rearrest risk, which corresponds to a negative sign in the parametric models, indicating a more rapid time until rearrest. Put somewhat differently, a higher hazard rate implies a shorter rearrest time, indicated by the opposite signs of the same coefficient in the proportional hazards and the parametric models.

To estimate a proportional hazards model, at least one risk variable must be included. For this reason, the unconditional risk-variable model (i.e., no risk variables included), designated by the tabular column "0," does not apply and is, therefore, not filled in. Also, because the proportional hazards model is semiparametric in the way described above, the tabular rows designating the "shape" and "scale" parameters do not apply and, likewise, are not filled in. See notes c and f for further discussion of the different risk-variable models which were estimated and of some special features of the shape and scale parameters.

b. We estimated the extreme value parameterization of this model. For details of the computational procedure, see D. Steinberg and P. Colla, SURVIVAL: A Supplementary Module for SYSTAT (Evanston, IL: SYSTAT, Inc, 1988). The extreme value parameterization yields coefficient estimates which are identical in magnitude and sign to those estimated by the nonextreme value parameterizations. Only the shape and scale parameters differ across parameterizations, and these are easily converted to one another. Technical estimation issues mainly influenced the decision to use this parameterization.

c. 0: The unconditional model (i.e., no risk variables included);
L: The legally-permissible risk-variable model;
L+R: The legally-permissible-plus-race risk-variable model;
A: The all (i.e., full) risk-variable model.

d. The number of risk variables in the model. This number sometimes changed across arrest transitions. A risk variable was included in the model at a particular arrest transition based on its distributional features. First, and most obviously, risk variables were included only when they could produce reliable estimates. For this reason, for example, the birth cohort subject's race during the juvenile period was included only at the first arrest transition; too few whites appeared at the later arrest transitions. Second, and less obviously, some variables were included in the model because of the way those variables were technically defined. For instance, the risk variable indicating the presence of an arrest for a prior UCR index crime was included only at the first arrest transition because the variable could take on different values only at this arrest transition. (The birth cohort subject might or might not have been previously arrested for a UCR property index crime.) At all subsequent transitions, the birth cohort subject must have had a prior arrest for a UCR index crime, at a minimum, the first arrest for a serious violent crime, which placed the subject in the study sample.

e. This value is minus two times the model's loglikelihood statistic, a statistic measuring how well the model matched the observed rearrest-time data. We have calculated minus two times this statistic because this value can be used to judge the comparative merits of nested risk-variable models, that is, of models whose coefficients are related in the form of superset to subset. One can compare, then, the "0" model to the "L," "L+R," and "A" models, the "L" model to the "L+R" and "A" models, and the "L+R" to the "A" model. These comparisons are transitive: if the "L" model is significantly superior to the "0" model, the "L+R" and the "A" models are also statistically superior because they include the "L" model.

When statistically comparing two risk-variable models, we performed the following steps: (1) calculated the difference between -2 times the loglikelihood of each model, which we have presented in the table, (2) calculated the degrees of freedom of the statistical test, which is the difference between the numbers of risk variables in the two models, (3) turned to a chi-square table, found the tabular cell entry which was at the intersection of the degrees-of-freedom and selected significance-level, and determined whether the value calculated in step 1 was greater than the tabular cell entry, and (4) reported that the comparison was statistically significant if the value calculated in step 1 was greater than the cell entry or, conversely, that the comparison was statistically nonsignificant if the opposite was true. An asterisk ("*") appearing after a loglikelihood value indicates that the broader risk-variable model (i.e., the superset) significantly improved (p. val. < .05) the explanatory capacity relative to the narrower risk-variable model (i.e., the subset) immediately adjacent to the left.

In some instances, the risk-variable model failed to converge. In such cases, it was impermissible to draw formal inferences about the magnitudes and signs of the model's coefficients. However, one can still loosely use the loglikelihood statistic of a nonconvergent model to compare the explanatory capacity of that model to other, related risk-variable models. We did not, however, formally compare a risk-variable model which failed to converge to any other model.

NC: The model did not converge.

f. The distribution's shape parameter defines the curvature of the hazard function. For some distributions, the shape parameter invariably produces a specific type of overall curvature, although the shape parameter's magnitude effects the details of that curvature. (For instance, the lognormal distribution is always single peaked, but the shape parameter governs the rate of incline to the peak and, in turn, the rate of decline thereafter.) For

Table 3.24 (cont.)

other distributions, the shape parameter produces generically different curvatures depending upon the magnitude of the parameter. The overall curvatures of the hazard functions of the Weibull and loglogistic distributions depend upon the magnitudes their shape parameters as follows:

Distribution	Value of Shape Parameter
Weibull (extreme value)	= 1, constant hazard function (i.e., the Weibull reduces to the exponential form), > 1, decreasing hazard function, < 1, increasing hazard function.
loglogistic	> 1, decreasing hazard function, ≤ 1, single-peaked hazard function.

An asterisk ("*") after the shape parameter indicates that the parameter was significantly different (p. val. < .05) from 1, in the indicated direction.

g. The "bullet" (".") before the variable name indicates the general name of the risk variable; the dash before the variable name indicates a specific level of the general risk variable.

h. [REF]: The omitted reference category of the categorical variable. The effect of a specific category of a categorical variable (e.g., the presence of a firearm) on the timing of rearrest is obtained by comparing the coefficient calculated for that category to the reference category (e.g., the absence of a weapon).

i. Only those coefficients are presented which were significant at p. val. < .05.

j. NE: The model was not estimated because there were too few cases to produce reliable results.

k. NS: Not significant at p. val < .05.

l. NA: Not applicable. The risk variable or risk-variable model did not apply at the arrest transition.

Table 3.25

1958 Birth Cohort: Adult Arrests for Violent Crimes--
Failure-Time Regression-Model Loglikelihoods, Shape and
Scale Parameters, and Significant Risk Variables by
Arrest Transition, Race, Parametric Distribution,
and Risk Variable Model; Combined Juvenile and
Young Adult Prior Criminal Records
(Construction Sample)

Panel A: 1st Arrest Transition--Total, Blacks, Whites
Panel B: 2nd Arrest Transition--Total, Blacks, Whites
Panel C: 3rd Arrest Transiti. -Blacks
Panel D: 4th Arrest Transition--Blacks
Panel E: 5th Arrest Transition--Blacks

Panel A.1: 1st Arrest Transition--Total (N = 911)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	22	0	4	5	22	0	4	5	22	0	4	5	22	0	4	5	22
-2 Loglikelihood[e]	NA	4,182	4,162*	4,078*	6,062	6,038*	6,014*	5,942	5,890	5,868*	5,848*	5,770*	5,886	5,866*	5,846*	5,760*	5,894	5,876*	5,856*	5,768*
Shape[f]	NA	--	--	--	1.0	1.0	1.0	1.0	1.8*	1.8*	1.8*	1.8*	1.6*	1.6*	1.6*	1.5*	3.0	3.0	3.0	2.9
Scale[f]	NA	--	--	--	8.3	8.6	9.0	7.5	9.0	9.4	10.0	5.0	8.5	8.8	9.5	3.1	8.7	9.1	9.8	3.1
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	.4	.2	NS	NA	-.4	-.2	NS	NA	-.7	- .4	NS	NA	-.7	- .5	NS	NA	-.7	- .5	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Other Weapon	NA	-.4	-.4	- .4	NA	.4	.2	.4	NA	.7	.7	.7	NA	.8	.7	.7	NA	.8	.8	.7
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	.7	.5	NA	NA	-.7	-.6	NA	NA	-1.2	-1.0	NA	NA	-1.2	-1.0	NA	NA	-1.3	-1.1
.Age at Arrest for Present Violent Crime	NA	NA	NA	>-.1	NA	NA	NA	NS	NA	NA	NA	< .1	NA	NA	NA	< .1	NA	NA	NA	< .1

Table 3.25--Panel A.1 (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
.Age at First Arrest	NA	NA	NA	>-.1	NA	NA	NA	< .1	NA	NA	NA	< .1	NA	NA	NA	< .1	NA	NA	NA	< .1
Prior Arrests for UCR Index Crimes																				
.Number (Log)	NA	NA	NA	NS	NA	NA	NA	- .3	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
First Prior UCR Index Crime																				
.Age	NA	NA	NA	< .1	NA	NA	NA	>-.1	NA	NA	NA	>-.1	NA	NA	NA	>-.1	NA	NA	NA	>-.1
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Other Weapon	NA	NA	NA	NS	NA	NA	NA	- .3	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.25 (cont.)

Panel A.2: 1st Arrest Transition--Blacks (N = 693)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21
-2 Loglikelihood[e]	NA	3,402	NA	3,336*	5,070	5,062	NA	5,002*	4,928	4,920	NA	4,856*	4,926	4,918	NA	4,850*	4,936	4,928	NA	4,858*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.8*	1.8*	NA	1.8*	1.6*	1.6*	NA	1.5*	2.9	2.9	NA	2.8
Scale[f]	NA	--	NA	--	8.2	8.0	NA	6.7	8.7	8.4	NA	3.6	8.1	7.9	NA	2.0	8.3	8.0	NA	1.2
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	-.4	NA	-.4	NA	.4	NA	.4	NA	NS	NA	.7	NA	NS	NA	.7	NA	.8	NA	.8
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Age at Arrest for Present Violent Crime	NA	NA	NA	>-.1	NA	NA	NA	NS	NA	NA	NA	<.1	NA	NA	NA	<.1	NA	NA	NA	<.1
.Age at First Arrest	NA	NA	NA	>-.1	NA	NA	NA	<.1	NA	NA	NA	<.1	NA	NA	NA	<.1	NA	NA	NA	<.1
Prior Arrests for UCR Index Crimes																				
.Number (Log)	NA	NA	NA	NS	NA	NA	NA	-.4	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
First Prior UCR Index Crime																				
.Age	NA	NA	NA	< .1	NA	NA	NA	>-.1	NA	NA	NA	>-.1	NA	NA	NA	>-.1	NA	NA	NA	>-.1

Table 3.25--Panel A.2 (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
Most Recent Prior UCR Index [illegible]e																				
.Age	NA	NA	NA	NS	NA	NA	NA	< .1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Other Weapon	NA	NA	NA	NS	NA	NA	NA	- .3	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.25 (cont.)

Panel A.3: 1st Arrest Transition--Whites (N = 218)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21	0	4	NA	21
-2 Loglikelihood[e]	NA	480	NA	450*	956	942*	NA	912*	934	920*	NA	890*	932	920*	NA	892*	932	920*	NA	888*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.8*	1.8*	NA	1.7*	1.7*	1.6*	NA	1.6*	3.3	3.2	NA	2.8
Scale[f]	NA	--	NA	--	9.0	9.7	NA	11.5	10.1	11.3	NA	11.3	9.7	10.8	NA	9.7	10.1	11.3	NA	13.3
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	.9	NA	.7	NA	-.9	NA	-.7	NA	-1.6	NA	NS	NA	-1.6	NA	-1.3	NA	-1.6	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.25 (cont.)

Panel B.1: 2nd Arrest Transition--Total (N = 325)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	5	26	0	4	5	26	0	4	5	26	0	4	5	26	0	4	5	26
-2 Loglikelihood[e]	NA	1,662	1,662	1,624*	2,736	2,718*	2,718	2,678*	2,672	2,658*	2,658	2,622*	2,670	2,658*	2,658	2,622*	2,674	2,660*	2,660	2,622*
Shape[f]	NA	--	--	--	1.0	1.0	1.0	1.0	1.7*	1.6*	1.6*	1.6*	1.4*	1.4*	1.4*	1.3*	2.6	2.5	2.5	2.4
Scale[f]	NA	--	--	--	7.7	7.2	7.3	9.6	8.0	7.3	7.4	9.1	7.4	6.8	6.7	9.0	7.5	6.8	6.5	9.0
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	-.3	-.3	-.3	NA	.3	.3	.4	NA	.5	.5	.5	NA	.5	.5	.5	NA	.5	.5	.5
.Weapon Used																				
-Firearm	NA	NS	NS	.5	NA	-.4	-.4	-.5	NA	NS	NS	NS	NA	NS	NS	-.9	NA	NS	NS	-1.0
-Other Weapon	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS	NA	NS	NS	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NS	NS	NA	NA	NS	NS	NA	NA	NS	NS	NA	NA	NS	NS	NA	NA	NS	NS
.Age at Arrest for Present Violent Crime	NA	NA	NA	>-.1	NA	NA	NA	<.1	NA	NA	NA	<.1	NA	NA	NA	<.1	NA	NA	NA	< .1
Prior Arrests for UCR Index Crimes																				
.Number (Log)	NA	NA	NA	NS	NA	NA	NA	- .5	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
.Mean Seriousness	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-1.4
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Any Priors																				
-Yes	NA	NA	NA	NS	NA	NA	NA	-1.6	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-2.7
-Unknown	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-No [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.25--Panel B.1 (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
.Mean Seriousness																				
-Known Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	.4	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Unknown Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
First Prior UCR Index Crime																				
.Type																				
-Robbery	NA	NA	NA	.5	NA	NA	NA	- .6	NA	NA	NA	- .9	NA	NA	NA	- .9	NA	NA	NA	-1.0
-Assault	NA	NA	NA	.6	NA	NA	NA	- .8	NA	NA	NA	-1.1	NA	NA	NA	-1.1	NA	NA	NA	-1.2
-Property [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Socioeconomic Status ≤ 15th Percentile	NA	NA	NA	.4	NA	NA	NA	- .6	NA	NA	NA	- .7	NA	NA	NA	- .6	NA	NA	NA	- .7

Table 3.25 (cont.)

Panel B.2: 2nd Arrest Transition--Blacks (N = 277)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	0[c]	L[c]	L+R[c]	A[c]	0	L	L+R	A	0	L	L+R	A	0	L	L+R	A	0	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25
-2 Loglikelihood[e]	NA	1,402	NA	1,370*	2,382	2,370*	NA	2,334*	2,346	2,336*	NA	2,302*	2,342	2,336	NA	2,302*	2,348	2,338*	NA	2,304*
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.5*	1.5*	NA	1.5*	1.3*	1.2*	NA	1.2*	2.3	2.3	NA	2.2
Scale[f]	NA	--	NA	--	7.7	7.4	NA	10.9	7.9	7.5	NA	10.9	7.3	7.1	NA	11.0	7.4	7.1	NA	11.2
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	.2	NA	.3	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	-.4	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Prior Status Offense	NA	NA	NA	NS	NA	NA	NA	.6	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
.Age at Arrest for Present Violent Crime	NA	NA	NA	>-.1	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
Prior Arrests for UCR Index Crimes																				
.Number (Log)	NA	NA	NA	NS	NA	NA	NA	-.6	NA	NA	NA	NS	NA	NA	NA	- .8	NA	NA	NA	- .8
.Mean Seriousness	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-1.2	NA	NA	NA	-1.5

Table 3.25--Panel B.2 (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Any Priors																				
-Yes	NA	NA	NA	1.7	NA	NA	NA	-1.9	NA	NA	NA	-2.6	NA	NA	NA	-2.6	NA	NA	NA	-3.0
-Unknown	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-No [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Mean Seriousness																				
-Known Adjudicated/Convicted	NA	NA	NA	-.5	NA	NA	NA	.6	NA	NA	NA	.7	NA	NA	NA	.8	NA	NA	NA	.9
-Unknown Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
First Prior UCR Index Crime																				
.Type																				
-Robbery	NA	NA	NA	.6	NA	NA	NA	-.7	NA	NA	NA	-.9	NA	NA	NA	-.9	NA	NA	NA	-1.0
-Assault	NA	NA	NA	NS	NA	NA	NA	-.7	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-Property [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Socioeconomic Status ≤ 15th Percentile	NA	NA	NA	.4	NA	NA	NA	-.5	NA	NA	NA	-.6	NA	NA	NA	NS	NA	NA	NA	-.6

Table 3.25 (cont.)

Panel B.3: 2nd Arrest Transition--Whites (N = 48)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE
-2 Loglikelihood[e]	NA	134	NA	NE	352	334	NA	NE	318	308*	NA	NE	316	306*	NA	NE	316	306*	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	2.7*	2.4*	NA	NE	2.3*	1.9*	NA	NE	4.0	3.3	NA	NE
Scale[f]	NA	--	NA	NE	7.8	5.5	NA	NE	8.8	5.1	NA	NE	7.9	4.0	NA	NE	8.0	4.0	NA	NE
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	-.9	NA	NE	NA	1.3	NA	NE	NA	2.2	NA	NE	NA	2.3	NA	NE	NA	2.3	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	-1.5	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	-1.0	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.25 (cont.)

Panel C: 3rd Arrest Transition--Blacks (N = 137)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25
-2 Loglikelihood[e]	NA	612	NA	596	1,172	1,172	NA	1,156	1,160	1,160	NA	1,146	1,160	1,158	NA	1,142	1,164	1,162	NA	1,142
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.4*	1.4*	NA	1.4*	1.2*	1.2*	NA	1.1	2.2	2.2	NA	2.0
Scale[f]	NA	--	NA	--	7.5	7.4	NA	5.2	7.7	7.6	NA	3.7	7.1	7.0	NA	2.6	7.2	7.1	NA	2.5
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.25 (cont.)

Panel D: 4th Arrest Transition--Blacks (N = 69)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25	0	4	NA	25
-2 Loglikelihood[e]	NA	254	NA	220*	560	556	NA	508*	542	538	NA	502*	542	538	NA	504*	542	538	NA	504
Shape[f]	NA	--	NA	--	1.0	1.0	NA	1.0	1.8*	1.8*	NA	1.4*	1.5*	1.5*	NA	1.2	2.8	2.7	NA	2.0
Scale[f]	NA	--	NA	--	7.2	6.8	NA	3.7	7.6	6.8	NA	.3	7.0	6.4*	NA	-2.9	7.0	6.6	NA	-4.6
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
.Weapon Used																				
-Firearm	NA	NS	NA	-1.5	NA	NS	NA	-1.4	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-Other Weapon	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS	NA	NS
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
.Incarcerated for a Prior UCR Index Crime	NA	NA	NA	NS	NA	NA	NA	1.4	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
Adjudicated/ Convicted for Prior UCR Index Crimes																				
.Mean Seriousness																				
-Known Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	2.0	NA	NA	NA	NS
-Unknown Adjudicated/Convicted	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS

Table 3.25--Panel D (cont.)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
Most Recent Prior UCR Index Crime																				
.Type																				
-Robbery	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-2.3
-Assault	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	-2.3
-Property [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Prior Arrest Involving a Weapon																				
-Firearm	NA	NA	NA	NS	NA	NA	NA	1.9	NA	NA	NA	2.1	NA	NA	NA	1.7	NA	NA	NA	1.7
-Other Weapon	NA	NA	NA	NS	NA	NA	NA	-1.5	NA	NA	NA	NS	NA	NA	NA	NS	NA	NA	NA	NS
-None [REF]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3.25 (cont.)

Panel E: 5th Arrest Transition--Blacks (N = 34)

Distribution Features, Statistics, and Risk Variables	Parametric Distribution and Risk Variable Model																			
	Proportional Hazards[a]				Exponential[b]				Weibull[b]				Loglogistic				Lognormal			
	O[c]	L[c]	L+R[c]	A[c]	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A	O	L	L+R	A
No. of Risk Variables[d]	NA	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE	0	4	NA	NE
-2 Loglikelihood[e]	NA	104	NA	NE	290	286	NA	NE	274	272	NA	NE	274	270	NA	NE	274	270	NA	NE
Shape[f]	NA	--	NA	NE	1.0	1.0	NA	NE	2.0*	2.0*	NA	NE	1.7*	1.6*	NA	NE	3.1	2.8	NA	NE
Scale[f]	NA	--	NA	NE	7.0	6.1	NA	NE	7.4	6.3	NA	NE	6.6	6.2	NA	NE	6.7	6.3	NA	NE
I. Permissible																				
Present Arrest for a Violent Crime																				
.Type[g]																				
-Robbery[g]	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Assault [REF][h]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
.Seriousness (Log)	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
.Weapon Used																				
-Firearm	NA	NS	NA	NE	NA	-.9	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-Other Weapon	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE	NA	NS	NA	NE
-None [REF][i]	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
II. Less Permissible and Impermissible																				
.Race	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

a. At each arrest transition, a proportional hazards (i.e., Cox) regression model was estimated in addition to the four parametric failure time regression models. This model permits one to estimate the effects of risk variables on the time-specific rearrest risk (i.e., the hazard rate) but, because it makes no assumption about the shape of the underlying parametric distribution which generated these risks, does not permit one to estimate distributional parameters. (Thus, it is a semiparametric model--parametric only in the coefficients.) The proportional hazards model assumes that the hazard functions of different levels, or strata, of a risk variable (e.g., low SES and high SES) are proportional to one another across time. For instance, low SES subjects might have a hazard rate that is twice as high as high SES subjects regardless of the point in time at which one contrasts the groups. While the proportional hazards model might have limited use with respect to js policy making which aims to devise interventions which are linked to time, because the model does not assign a specific time-varying shape to the hazard function, it does have an important feature: it is robust. Risk-variable coefficients are reliably estimated by the proportional hazards model across a variety of parametric distributions (e.g., exponential, Weibull, loglogistic). Because it is robust, the proportional hazards model can be used as a baseline model against which to compare the coefficient estimates produced by the other parametric models. We use the proportional hazards model, therefore, to check the consistency and, in turn, plausibility of our results.

As the name hints, the risk variables in the proportional hazards model influence the <u>hazard function</u>, which is the rearrest function in this study. A positive coefficient indicates that the presence of a risk characteristic increases the rearrest function; a negative coefficient indicates the reverse. To compare the effect of a coefficient estimated by the proportional hazards model to the effect of a coefficient estimated by one of the other parametric models, one simply flips the sign of the proportional hazards coefficient. This is done because the parametric models produce coefficient estimates based on the relationship of the risk variable to the timing of rearrest, not to the rearrest (i.e., hazard) rate. When a coefficient in the proportional

hazards model has a positive sign, the risk variable increases the rearrest risk, which corresponds to a negative sign in the parametric models, indicating a more rapid time until rearrest. Put somewhat differently, a higher hazard rate implies a shorter rearrest time, indicated by the opposite signs of the same coefficient in the proportional hazards and the parametric models.

To estimate a proportional hazards model, at least one risk variable must be included. For this reason, the unconditional risk-variable model (i.e., no risk variables included), designated by the tabular column "0," does not apply and is, therefore, not filled in. Also, because the proportional hazards model is semiparametric in the way described above, the tabular rows designating the "shape" and "scale" parameters do not apply and, likewise, are not filled in. See notes c and f for further discussion of the different risk-variable models which were estimated and of some special features of the shape and scale parameters.

b. We estimated the extreme value parameterization of this model. For details of the computational procedure, see D. Steinberg and P. Colla, SURVIVAL: A Supplementary Module for SYSTAT (Evanston, IL: SYSTAT, Inc, 1988). The extreme value parameterization yields coefficient estimates which are identical in magnitude and sign to those estimated by the nonextreme value parameterizations. Only the shape and scale parameters differ across parameterizations, and these are easily converted to one another. Technical estimation issues mainly influenced the decision to use this parameterization.

c. 0: The unconditional model (i.e., no risk variables included);
L: The legally-permissible risk-variable model;
L+R: The legally-permissible-plus-race risk-variable model;
A: The all (i.e., full) risk-variable model.

d. The number of risk variables in the model. This number sometimes changed across arrest transitions. A risk variable was included in the model at a particular arrest transition based on its distributional features. First, and most obviously, risk variables were included only when they could produce reliable estimates. For this reason, for example, the birth cohort subject's race during the juvenile period was included only at the first arrest transition; too few whites appeared at the later arrest transitions. Second, and less obviously, some variables were included in the model because of the way those variables were technically defined. For instance, the risk variable indicating the presence of an arrest for a prior UCR index crime was included only at the first arrest transition because the variable could take on different values only at this arrest transition. (The birth cohort subject might or might not have been previously arrested for a UCR property index crime.) At all subsequent transitions, the birth cohort subject must have had a prior arrest for a UCR index crime, at a minimum, the first arrest for a serious violent crime, which placed the subject in the study sample.

e. This value is minus two times the model's loglikelihood statistic, a statistic measuring how well the model matched the observed rearrest-time data. We have calculated minus two times this statistic because this value can be used to judge the comparative merits of nested risk-variable models, that is, of models whose coefficients are related in the form of superset to subset. One can compare, then, the "0" model to the "L," "L+R," and "A" models, the "L" model to the "L+R" and "A" models, and "L+R" to the "A" model. These comparisons are transitive: if the "L" model is significantly superior to the "0" model, the "L+R" and the "A" models are also statistically superior because they include the "L" model.

When statistically comparing two risk-variable models, we performed the following steps: (1) calculated the difference between -2 times the loglikelihood of each model, which we have presented in the table, (2) calculated the degrees of freedom of the statistical test, which is the difference between the numbers of risk variables in the two models, (3) turned to a chi-square table, found the tabular cell entry which was at the intersection of the degrees-of-freedom and selected significance-level, and determined whether the value calculated in step 1 was greater than the tabular cell entry, and (4) reported that the comparison was statistically significant if the value calculated in step 1 was greater than the cell entry or, conversely, that the comparison was statistically nonsignificant if the opposite was true. An asterisk ("*") appearing after a loglikelihood value indicates that the broader risk-variable model (i.e., the superset) significantly improved (p. val. < .05) the explanatory capacity relative to the narrower risk-variable model (i.e., the subset) immediately adjacent to the left.

In some instances, the risk-variable model failed to converge. In such cases, it was impermissible to draw formal inferences about the magnitudes and signs of the model's coefficients. However, one can still loosely use the loglikelihood statistic of a nonconvergent model to compare the explanatory capacity of that model to other, related risk-variable models. We did not, however, formally compare a risk-variable model which failed to converge to any other model.

NC: The model did not converge.

f. The distribution's shape parameter defines the curvature of the hazard function. For some distributions, the shape parameter invariably produces a specific type of overall curvature, although the shape parameter's magnitude effects the details of that curvature. (For instance, the lognormal distribution is always single peaked, but the shape parameter governs the rate of incline to the peak and, in turn, the rate of decline thereafter.) For

Table 3.25 (cont.)

other distributions, the shape parameter produces generically different curvatures depending upon the magnitude of the parameter. The overall curvatures of the hazard functions of the Weibull and loglogistic distributions depend upon the magnitudes their shape parameters as follows:

Distribution	Value of Shape Parameter
Weibull (extreme value)	= 1, constant hazard function (i.e., the Weibull reduces to the exponential form), > 1, decreasing hazard function, < 1, increasing hazard function.
loglogistic	> 1, decreasing hazard function, ≤ 1, single-peaked hazard function.

An asterisk ("*") after the shape parameter indicates that the parameter was significantly different (p. val. < .05) from 1, in the indicated direction.

g. The "bullet" (".") before the variable name indicates the general name of the risk variable; the dash before the variable name indicates a specific level of the general risk variable.

h. [REF]: The omitted reference category of the categorical variable. The effect of a specific category of a categorical variable (e.g., the presence of a firearm) on the timing of rearrest is obtained by comparing the coefficient calculated for that category to the reference category (e.g., the absence of a weapon).

i. Only those coefficients are presented which were significant at p. val. < .05.

j. NE: The model was not estimated because there were too few cases to produce reliable results.

k. NS: Not significant at p. val < .05.

l. NA: Not applicable. The risk variable or risk-variable model did not apply at the arrest transition.

Figure 2.1

Definitions and Scaling of Risk Variables

Set 1: <u>Legally/Ethically Permissible Risk Variables--A Strict Just Deserts Interpretation</u>

1. Type of Present Arrest for a Violent Crime

   * Robbery (0 = assault, includes homicide, forcible rape, aggravated assault; 1 = robbery)

2. Seriousness of the Present Arrest for a Violent Crime (log)

3. Weapon Used in the Present Arrest for a Violent Crime

   * Firearm (0 = other; 1 = firearm)
   * Other Weapon (0 = other; 1 = other weapon)
   * No Weapon (0 = other; 1 = no weapon) [REFERENCE CATEGORY][a]

Set II. <u>Legally/Ethically Less Permissible and Impermissible Risk Variables</u>

A. <u>Less Permissible--A Modified Just Deserts Interpretation</u>

4. Prior Arrest for a UCR Index Crime (0 = yes; 1 = no)[b]

5. Number of Prior Arrests for UCR Index Crimes (log)

6. Mean Seriousness of Prior Arrests for UCR Index Crimes (log)

7. Adjudicated/Convicted for a Prior UCR Index Crime

   * Any Prior Adjudication/Conviction (0 = no; 1 = yes)
   * Unknown Whether Any Prior Adjudication/Conviction (0 = known if any prior; 1 = unknown)
   * No Prior Adjudication/Conviction (0 = yes, unknown; 1 = no prior) [REFERENCE CATEGORY]

8. Mean Seriousness of Prior UCR Index Crimes

   * Known to have been Adjudicated/Convicted (log)
   * Unknown to have been Adjudicated/Convicted (log)

9. Prior Status Offense (0 = no; 1 = yes)

10. Prior Arrest Involving a Weapon

    * Firearm (0 = other; 1 = firearm)
    * Other Weapon (0 = other; 1 = other weapon)
    * No Weapon (0 = other; 1 = no weapon) [REFERENCE CATEGORY]

11. Type of First Prior Arrest for a UCR Index Crime

    * Robbery (0 = other; 1 = robbery)
    * Assault (0= other; 1 = homicide, forcible rape, aggravated assault)
    * Property (0 = other; 1 = burglary, larceny/theft, motor vehicle theft) [REFERENCE CATEGORY]

12. Seriousness of the First Prior Arrest for a UCR Index Crime (log)

Figure 2.1 (cont.)

13. Type of Most Recent Prior Arrest for a UCR Index Crime

    * Robbery (0 = other; 1 = robbery)
    * Assault (0 = other; 1 = homicide, forcible rape, aggravated assault)
    * Property (0 = other; 1 = burglary, larceny/theft, motor vehicle theft) [REFERENCE CATEGORY]

14. Seriousness of the Most Recent Prior Arrest for a UCR Index Crime (log)

15. Incarcerated for a Prior UCR Index Crime (0 = no; 1 = yes)

B. <u>Suspect Classes and Impermissible</u>

16. Race (0 = white; 1 = black)

17. Socioeconomic Status (0 = $\geq$ 15th percentile; 1 = < 15th percentile)

18. Age (in Months) at the Time of Arrest for the Present UCR Violent Index Crime

19. Age (in Months) at the Time of First Arrest

20. Age (in Months) at the Time of Arrest for the First Prior UCR Index Crime

21. Age (in Months) at the Time of Arrest for the Most Recent Prior UCR Index Crime

---

a. This is the reference, or comparison, category used for appraising the effect of a risk variable on the timing of rearrest. For example, the effects of the presence of a "firearm" or some "other weapon" on the timing of rearrest are separately compared to the effect of "no weapon."

b. At some arrest transitions, some birth cohort subjects did not have any prior arrests for UCR index crimes. This variable (coded this way) permitted us to analyze intelligibly the effects on rearrest timing of aspects of a subject's prior UCR-index-crime record (e.g., seriousness, age at first prior, age at most rescent prior) for those subjects who had such a record, while adjusting for those subjects who do not have such a record.

Figure 2.2

Failure Time Functions: Definitions and Computational Conventions[a]

1. Hazard Function

   a. Definition:

$$h(t) = \lim_{\Delta t \to 0} \frac{\text{Prob}\left\{\begin{array}{c}\text{An Arrested Violent Criminal Who Has Not Been}\\ \text{Rearrested by Time t Will be Rearrested in the}\\ \text{Time Interval } [t, t+\Delta]\end{array}\right\}}{\Delta t}$$

   b. Computational Convention:

$$\hat{h}(t) = \frac{\left\{\begin{array}{c}\text{Number of Arrested Violent Criminals Who Have Been}\\ \text{Rearrested in the Time Interval Beginning at Time t}\end{array}\right\}}{\left\{\begin{array}{c}\text{Number of Arrested Violent Criminals Who}\\ \text{Have Not Been Rearrested by Time t}\end{array}\right\}\left\{\begin{array}{c}\text{Width of the}\\ \text{Time Interval}\end{array}\right\}}$$

2. Survival Function

   a. Definition:

$$S(t) = \text{Prob}\left\{\begin{array}{c}\text{An Arrested Violent Criminal Will}\\ \text{Not Be Rearrested until after Time t}\end{array}\right\}$$

$$= \text{Prob } (T > t)$$

   b. Computational Convention:

$$\hat{S}(t) = \frac{\left\{\begin{array}{c}\text{Number of Arrested Violent Criminals}\\ \text{Who Have Been Rearrested after Time t}\end{array}\right\}}{(\text{Total Number of Arrested Violent Criminals})}$$

3. Probability Density Function

   a. Definition

$$f(t) = \lim_{\Delta t \to 0} \frac{\text{Prob}\left\{\begin{array}{c}\text{An Arrested Violent Criminal Will Be}\\ \text{Rearrested in Time Interval } [t, t+\Delta t]\end{array}\right\}}{\Delta t}$$

   b. Computational Convention:

$$\hat{f}(t) = \frac{\left\{\begin{array}{c}\text{Number of Arrested Violent Criminals Who Have Been}\\ \text{Rearrested in the Time Interval Beginning at Time t}\end{array}\right\}}{\left\{\begin{array}{c}\text{Total Number of}\\ \text{Arrested Violent Criminals}\end{array}\right\}\left\{\begin{array}{c}\text{Width of the}\\ \text{Time Interval}\end{array}\right\}}$$

---

a. The definitions and computational conventions apply in the absence of censored observations. Also, computational conventions pertain to nonparametric failure time procedures.

Figure 3.1

Juveniles:
Risk Variables Available for Use in the Failure Time Regression Models by Arrest Transition and Race

	Risk Variables	1st T[a]	1st B[a]	1st W[a]	2nd B	3rd B	4th B	5th B
Set I.	Permissible							
	Present Arrest for a Violent Crime							
	Type							
	- Robbery	X	X	X	X	X	X	X
	- Assault [REF][b]							
	Seriousness	X	X	X	X	X	X	X
	Weapon Used							
	- Firearm	X	X	X	X	X	X	X
	- Other Weapon	X	X	X	X	X	X	X
	- None [REF]							
Set II.	Less Permissible and Impermissible							
	Race[c]	X	-	-	-	-	-	-
	Prior Status Offense	X	X	X	X	X	X	X
	Age at Arrest for Present Violent Crime	X	X	X	X	X	X	X
	Age at First Arrest	X	X	X	X	X	X	X
	Prior Arrests for UCR Index Crimes							
	- Any Priors[d]	X	X	X	-	-	-	-
	- Number	X	X	X	X	X	X	X
	- Mean Seriousness	X	X	X	X	X	X	X
	Adjudicated/Convicted For Prior UCR Index Crimes							
	Any Priors							
	- Yes	X	X	X	X	X	X	X
	- Unknown	X	X	X	X	X	X	X
	- No [REF]							
	Mean Seriousness							
	- Known Adjudicated/ Convicted	X	X	X	X	X	X	X
	- Unknown Adjudicated/ Convicted	X	X	X	X	X	X	X
	Incarcerated for a Prior UCR Index Crime[e]	-	-	-	X	X	X	X

Figure 3.1 (cont.)

Risk Variables	1st T	1st B	1st W	2nd B	3rd B	4th B	5th B
First Prior UCR Index Crime							
Type							
- Robbery[f]	-	-	-	X	X	X	X
- Assault[f]	-	-	-	X	X	X	X
- Property [REF]							
Age	X	X	X	X	X	X	X
Seriousness	X	X	X	X	X	X	X
Most Recent Prior UCR Index Crime							
Type							
- Robbery[g]	-	-	-	X	X	X	X
- Assault[g]	-	-	-	X	X	X	X
- Property [REF]							
Age	X	X	X	X	X	X	X
Seriousness	X	X	X	X	X	X	X
Prior Arrest Involving a Weapon							
- Firearm	X	X	X	X	X	X	X
- Other Weapon	X	X	X	X	X	X	X
- None [REF]							
Socioeconomic Status ≤ 15th Percentile	X	X	X	X	X	X	X

a. T: Total
   B: Blacks
   W: Whites

b. The suppressed reference category.

c. The race variable was used only at the 1st arrest transition. There were too few whites at later arrest transitions to support a reliable analysis.

d. The variable could take on different values only at the 1st arrest transition. After this transition, all subjects had at least one prior arrest for a UCR index crime--their first arrest for a violent crime.

e. There were too few prior incarcerations at the 1st arrest transition.

f. This variable could be used only after the 1st arrest transition because at the 1st arrest transition the first prior UCR index crime could only be a property index crime.

g. See note e. The same explanation applies.

Figure 3.2

Adults:
Risk Variables Available for Use in the Failure Time
Regression Models by Arrest Transition
and Race

Risk Variables	1st T[a]	1st B[a]	1st W[a]	2nd T	2nd B	2nd W	3rd B	4th B	5th B
Set I. Permissible									
Present Arrest for a Violent Crime									
Type									
- Robbery	X	X	X	X	X	X	X	X	X
- Assault [REF][b]									
Seriousness	X	X	X	X	X	X	X	X	X
Weapon Used									
- Firearm	X	X	X	X	X	X	X	X	X
- Other Weapon	X	X	X	X	X	X	X	X	X
- None [REF]									
Set II. Less Permissible and Impermissible									
Race[c]	X	-	-	X	-	-	-	-	-
Prior Status Offense	-	-	-	-	-	-	-	-	-
Age at Arrest for Present Violent Crime	X	X	X	X	X	X	X	X	X
Age at First Arrest	X	X	X	X	X	X	X	X	X
Prior Arrests for UCR Index Crimes									
- Any Priors[d]	X	X	X	-	-	-	-	-	-
- Number	X	X	X	X	X	X	X	X	X
- Mean Seriousness	X	X	X	X	X	X	X	X	X
Adjudicated/Convicted For Prior UCR Index Crimes									
Any Priors									
- Yes	X	X	X	X	X	X	X	X	X
- Unknown	X	X	X	X	X	X	X	X	X
- No [REF]									
Mean Seriousness									
- Known Adjudicated/Convicted	X	X	X	X	X	X	X	X	X
- Unknown Adjudicated/Convicted	X	X	X	X	X	X	X	X	X
Incarcerated for a Prior UCR Index Crime[e]	-	-	-	X	X	X	X	X	X

Figure 3.2 (cont.)

Risk Variables	1st			2nd			3rd	4th	5th
	T[a]	B[a]	W[a]	T	B	W	B	B	B
First Prior UCR Index Crime									
Type									
- Robbery[f]	-	-	-	X	X	X	X	X	X
- Assault[f]	-	-	-	X	X	X	X	X	X
- Property [REF]									
Age	X	X	X	X	X	X	X	X	X
Seriousness	X	X	X	X	X	X	X	X	X
Most Recent Prior UCR Index Crime									
Type									
- Robbery[g]	-	-	-	X	X	X	X	X	X
- Assault[g]	-	-	-	X	X	X	X	X	X
- Property [REF]									
Age	X	X	X	X	X	X	X	X	X
Seriousness	X	X	X	X	X	X	X	X	X
Prior Arrest Involving a Weapon									
- Firearm	X	X	X	X	X	X	X	X	X
- Other Weapon	X	X	X	X	X	X	X	X	X
- None [REF]									
Socioeconomic Status ≤ 15th Percentile	X	X	X	X	X	X	X	X	X

a. T: Total
B: Blacks
W: Whites

b. The suppressed reference category.

c. The race variable was used only at the 1st and 2nd arrest transitions. There were too few whites at later arrest transitions to support a reliable analysis.

d. The variable could take on different values only at the 1st arrest transition. After this transition, all subjects had at least one prior arrest for a UCR index crime--their first arrest for a violent crime.

e. There were too few prior incarcerations at the 1st arrest transition.

f. This variable could be used only after the 1st arrest transition because at the 1st arrest transition the first prior UCR index crime could only be a property index crime.

g. See note e. The same explanation applies.