

LONG-TERM CRIME DESISTANCE AND RECIDIVISM PATTERNS—EVIDENCE FROM THE ESSEX COUNTY CONVICTED FELON STUDY*

MEGAN C. KURLYCHEK
SHAWN D. BUSHWAY

School of Criminal Justice
University at Albany

ROBERT BRAME

Department of Criminal Justice and Criminology
University of North Carolina at Charlotte

KEYWORDS: crime desistance, recidivism, survival models, criminal careers

Two conflicting definitions of desistance exist in the criminology literature. The first definition is instantaneous desistance in which an offender simply chooses to end a criminal career instantaneously moving to a zero rate of offending (Blumstein et al., 1986). The second definition views desistance as a process by which the offending rate declines steadily over time to zero or to a point close to zero (Bushway et al., 2001; Laub and Sampson, 2001; Leblanc and Loeber, 1998). In this article, we capitalize on the underlying assumptions of several parametric survival distributions to gain a better understanding of which of these models best describes actual patterns of desistance. All models are examined using 18 years of follow-up data on a cohort of felony convicts in Essex County, NJ. Our analysis leads us to three conclusions. First,

Additional supporting information can be found in the listing for this article in the Wiley Online Library at <http://onlinelibrary.wiley.com/doi/10.1111.crim2012.50.issue-1/issuetoc>. This article was accepted under the editorship of Denise Gottfredson.

* We wish to thank Kelly Socia, who helped us understand the Essex County data set, and seminar participants at the University at Albany and the Reentry Institute at John Jay College. We also would like to thank Alex Piquero and the anonymous reviewers whose thoughtful comments helped to improve the accuracy and quality of this article. All errors remain our own. Direct correspondence to Megan C. Kurlychek, School of Criminal Justice, University at Albany, 135 Western Avenue, Albany, NY 12222 (e-mail: mkurlychek@albany.edu).

some people have already desisted at the beginning of the follow-up period, which is consistent with the notion of “instantaneous desistance.” Second, a three-parameter model that allows for a turning point in the risk of recidivism followed by a long period of decline fits the data best. This conclusion suggests that for those offenders active at the start of the study period, the risk of recidivism is declining over time. However, we also find that a simpler two-group model fits the data almost as well and gains superiority in the later years of follow-up. This last point is particularly relevant as it suggests that the observed gradual decline in the hazard over time is a result of a compositional effect rather than of a pattern of individually declining hazards.

Two conflicting definitions of *desistance* may be found in the criminology literature. The first definition, which comes from the criminal career tradition, is instantaneous desistance. Constructed by Blumstein et al. (1986) in early criminal career research, this model provides a simple description of an offending career in which an offender 1) begins a criminal career, 2) offends at a constant rate while active, and 3) finally desists from offending, instantly moving to a zero rate of offending. This point of view is supported by qualitative descriptions of desistance that emphasize a discrete point in time when an individual decides to “go straight” (Baskin and Sommers, 1998; Maltz, 1996; Maruna, 1999) and by findings from the reentry literature suggesting that some offenders abruptly and permanently desist (National Research Council, 2008). This original model was later modified because of empirical evidence of higher than expected offending many years after the last event. This modification was described as “intermittency,” the idea that an offender may “restart” a criminal career after a period of inactivity (Barnett, Blumstein, and Farrington, 1989).

The second, and more recent, definition of *desistance* allows the offending rate to vary over a criminal career. More specifically, this model views desistance as a process by which the offending rate declines steadily over time to a stable point, which is either zero or close to zero (Bushway et al., 2001; Laub and Sampson, 2001; Leblanc and Loeber, 1998). This type of desistance path has been described as a “glide path” to zero and is consistent with the idea of life-course trajectories that exhibit both continuity and gradual behavioral change (Sampson and Laub, 2003).

Recently, criminologists have begun to use hazard models to examine the long-term hazards of recidivism, in so-called redemption research (Blumstein and Nakamura, 2009; Bushway, Nieuwbeerta, and Blokland, 2011; Kurlychek, Brame, and Bushway, 2006, 2007; Soothill and Francis, 2009).¹

1. The term *redemption* was first applied to this question by Blumstein and Nakamura (2009).

This largely atheoretical research attempts to determine whether individuals with criminal history records ever look like those without criminal history records in terms of future offending risk. The finding that individuals are in fact redeemed has been used to motivate public policy limiting the use of criminal history records in employment decisions (Bushway, Nieuwebeerta, and Blokland, 2011).

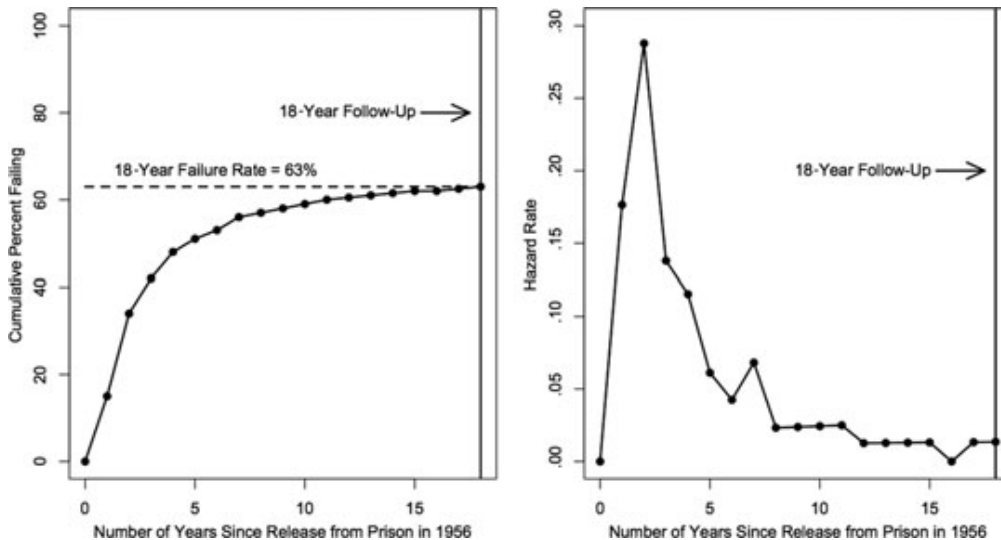
Each theoretical depiction of desistance has different implications. If individuals desist “instantaneously,” then waiting only adds information—it does not reduce risk. In this case, on the one hand, policy makers could improve social outcomes by finding ways to identify desisters earlier. On the other hand, if individuals are “gliding” to desistance, the process itself is important and the full time to redemption is needed to reach this low level of risk.

Different parametric survival models exist that fit the form of either a constant rate of offending followed by an instantaneous drop to zero or a gradual decline in offending rates over time (Maltz, 1996: 29–32). The approach of this article is to study the nature of desistance over the long term by identifying the model that best fits patterns observed in an 18-year follow-up study of convicted felony offenders in Essex County, NJ. If we find support for the idea that individuals “instantaneously” desist, we can build on earlier work by others such as Maltz (1984), Schmidt and Witte (1988), and Brame, Bushway, and Paternoster (2003) to provide a new estimate of the size of the population exhibiting an abrupt and permanent termination of criminal offending. Unlike this earlier work, which mostly relies on follow-up periods of a few years or less, our estimates are based on a much longer time horizon. In the next section, we will describe the current literature on recidivism, before describing the models, data, and results. In the final section, we consider the implications of our findings and suggest some directions for future research.

LITERATURE REVIEW

Once someone has been defined as an offender—regardless of whether an arrest, conviction, or incarceration experience is the basis for that definition—the recidivism clock begins to tick. Criminologists have long understood the importance of measuring how long this clock ticks before the next criminal event occurs (Maltz, 1996: 29–32). Measures of recidivism that document only whether recidivism has occurred do not distinguish between people who fail more quickly and those who fail more slowly. This distinction matters because people who fail more quickly also are failing at a higher rate over any well-defined time interval (Maltz, 1984; Schmidt and Witte, 1988).

Figure 1. Recidivism Patterns in the Federal Bureau of Prisons Sample of 1956 Releasees ($N = 903$)



Source: Adapted from information presented in Kitchener, Schmidt, and Glaser (1977: 10).

As studies documenting the distribution of recidivism waiting times have accumulated, a consensus has emerged that “time to recidivism distributions” have some features that appear with a good deal of regularity regardless of the type of offender studied, the actual measure of recidivism employed, or the length of the follow-up period. In particular, the most prominent feature of these studies is the positive skew in the distribution of waiting times (Beck and Shipley, 1989; Carr-Hill and Carr-Hill, 1972; Dean, Brame, and Piquero, 1996; DeJong, 1997; Kurlychek, Brame, and Bushway, 2006, 2007; Langan and Levin, 2002; Lattimore and Baker, 1993; Maltz, 1984, 1996; Schmidt and Witte, 1988; Stollmack and Harris, 1974; Visser, Lattimore, and Linster, 1991).

An early study by Kitchener, Schmidt, and Glaser (1977) is illustrative. This research project followed up 903 offenders released from federal prisons in 1956 for 18 years to determine what percentage had experienced a parole revocation or had been returned to prison for serious offenses (see also Schmidt and Witte, 1988: 68). An important feature of the study was its attention to how the percentage of cases recidivating accumulated over time.

As the left-hand side of figure 1 shows, the earliest years of the follow-up period were also the years where the cumulative percentage increased most steeply. The right-hand side of figure 1 presents a plot of the hazard rate from the 1956 Bureau of Prisons data—a key estimand in any recidivism study. It is obtained by dividing the number of people who fail in a given

time period by the number of people who have not yet failed by the beginning of that time period.

Two important characteristics of the hazard rate plot are shown in figure 1. The first is that a clear turning point in the risk of recidivism occurs at the 2-year mark. This turning point means that the period of greatest risk is close to, but not immediately after, the beginning of the follow-up period. The second key feature of the hazard rate plot is a rapid decline from the peak at 2 years followed by a long, gradually declining right-hand tail. Taken together, these results point to a general reduction in the risk of recidivism over time. The general tendency for recidivism risk to decline over time is among the best replicated results in empirical criminology. It is probably not an exaggeration to say that any recidivism study with more than a 2- or 3-year follow-up period that did not find a downward-sloping marginal hazard would be immediately suspect.

This sort of pattern has been observed in other fields of study ranging from automobile accidents, to insurance claims, to length-of-unemployment spells, to loan defaults, to marriage dissolution, to death rates among cancer patients (Klugman, Panjer, and Wilmot, 2008; Kurlychek, Brame, and Bushway, 2006, 2007). Yet, as consistent as this pattern has been in criminology and elsewhere, it actually raises as many questions as it answers. Knowing that a waiting-time distribution is positively skewed does not in and of itself help us understand why it is skewed.

It might be tempting to adopt the obvious interpretation: As the waiting time increases, each individual's propensity to fail declines. Hence, a skewed waiting-time distribution simply reflects the fact that individual-level risk declines in the same skewed way over time. Such an interpretation is plausible. As people who have offended in the past gain experience, skills, and abilities in living their lives without crime, individual risk could very well exhibit a meaningful decline. This "experiential" effect would be supportive of the glide-path notion of desistance as a gradual process. That is, an individual's risk of recidivism does actually decline over time as predicted by Laub and Sampson (2001) among others. But such an explanation also could run afoul of the ecological fallacy: an untested assumption that individual-level risk trends track those of the overall population risk trend (Maltz, 1996: 32).

In fact, a compelling rival explanation for commonly observed patterns in recidivism waiting-time distributions is that 1) individual-level risk is fixed and time-stable at the beginning of the follow-up period; 2) the sample of offenders being studied comprises low-, medium-, and high-risk individuals; and 3) high-risk individuals fail most quickly, whereas medium-risk individuals fail more slowly, and low-risk individuals fail still more slowly or not at all (Schmidt and Witte, 1988: 69). Put more simply, the high rate of offending observed immediately after the original event is a result of the high-risk people failing quickly. The remaining sample in the risk set

then comprises the medium- and low-risk individuals who would have taken longer to fail regardless. In the extreme case, the low-risk individuals are “instantaneous” desisters as described by the original criminal career researchers (Barnett, Blumstein, and Farrington, 1989; Blumstein et al. 1986; Blumstein, Farrington, and Moitra, 1985; Blumstein and Moitra, 1980; Maltz, 1996). For many criminologists, this “compositional change” hypothesis is just as compelling as the “experiential decline” hypothesis. But, at least to date, criminologists have put forth little effort to use recidivism waiting-time distributions to distinguish between these two explanations. We now turn to a subset of the literature that has made some limited progress in this area.

SPLIT-POPULATION MODELS

It is difficult to differentiate empirically between the compositional (e.g., the characteristics of the sample change over time) and the experiential (e.g., the individuals in the sample have dynamic characteristics that change over time) explanations for the declining hazard rate. Both explanations predict that recidivism risk will decline over time. However, an interesting special case of the “compositional change” framework arises when a group of “instantaneous” desisters abruptly and permanently terminates offending activity at the outset of the waiting period. These offenders were first formally incorporated into the study of recidivism by Blumstein and Larson (1969: 219–25) in their discussion of total criminal justice system models.

However, the study of recidivism and desistance among individual offenders has always been challenging because studies must end at some point. Today’s apparent desister can always become tomorrow’s repeat offender. It is not surprising to discover that criminologists simultaneously began to turn their attention to the problem of “censoring” that inevitably occurs in recidivism studies. In brief, the censoring problem is caused by the limited follow-up period in which any recidivism study must be conducted. During this time, some people are observed to fail, whereas others “survive” without a failure to the end of the follow-up period. All we know about the survivors is that they did not recidivate by the end of the follow-up period—not whether they ever failed. Parametric survival-time models characterize these individuals as “censored.” All other things equal, shorter follow-up periods correspond to higher rates of censoring (Carr-Hill and Carr-Hill, 1972; Maltz, 1984: 71–5; Maltz and McCleary, 1977; Stollmack and Harris, 1974; Witte and Schmidt, 1977).

Here is a closely related problem: If we could eliminate the censoring problem and follow offenders for the rest of their lives, would the limiting failure rate approach 100 percent? Building on this work and the earlier work of Anscombe (1961), Maltz and McCleary (1977), and Maltz (1984:

79–81), Schmidt and Witte (1984: 92–4; 1988: 66–9) noted that an important feature of standard parametric survival-time models—the type that often are used to model recidivism waiting-time distributions—is an assumption that if the follow-up period were lengthened to an infinitely long time, the cumulative percentage of individuals who fail would approach 100 percent. Although such approaches are clearly reasonable in applications such as waiting times for machine and bridge failure, they are not reasonable for many social outcomes such as criminal recidivism (Maltz, 1996: 29). As Schmidt and Witte (1988: 69) pointed out, such models are doomed to generate estimates of long-term recidivism rates that are too high.

The innovation of the split-population framework is its allowance for the possibility that some fraction of the offender population has abruptly and permanently desisted by the beginning of the follow-up period for a recidivism study. The complement of this population comprises people who are still active offenders at the beginning of the follow-up period. Note that an individual can be an active offender and still not be observed to fail within any finite follow-up period. So, split-population models can be decomposed into two distinct parts: 1) an analysis that estimates the proportion of people who have abruptly and permanently desisted by the beginning of the follow-up period and 2) an analysis of the waiting time to recidivism distribution for the subset of individuals who are still active at the beginning of the follow-up period. This decomposition guarantees that as the follow-up period grows infinitely long, the limiting failure rate will be the proportion of people who are active offenders (in general, a number that is significantly less than one). And, as Maltz (1996: 30) noted, the split-population framework closely approximates the distinctions among “innocents,” “persisters,” and “desisters” emphasized by Blumstein and Moitra (1980), Blumstein, Farrington, and Moitra (1985), and Barnett, Blumstein, and Farrington (1987, 1989).

Several recidivism studies have used split-population models to study the phenomenon of abrupt and permanent desistance. The most prominent contributions to this literature are Schmidt and Witte (1988) and Maltz (1984). Schmidt and Witte’s (1988) study of two North Carolina prison release cohorts (fiscal years 1978 and 1980) showed that split-population models offered dramatic improvements in predictions about recidivism. By defining recidivism as a return to prison within a 70–81-month follow-up period (for people released from prison in 1978), and by estimating a model allowing for an exponential (time-constant) hazard rate among the subset of active offenders, they found that the prevalence of abrupt and permanent desistance was on the order of 59 percent (Schmidt and Witte, 1988: 70). For the 1980 cohort and a 46–57-month follow-up period, the proportion of abrupt and permanent desisters was estimated to be approximately 54 percent.

Maltz (1984) discussed several early efforts to estimate split-population models. He then presented split-population models with the exponential distribution for recidivism waiting times among the active offenders using several different data sets with relatively short follow-up periods (up to 6 years long). For populations with heterogeneous groups of offenders, Maltz found consistent evidence that the split-population model performed better than the alternatives.² Other, more recent, applications of split-population models in criminology such as those by Lattimore and Baker (1993), Dean, Brame, and Piquero (1996), and DeJong (1997) suggest that distinguishing between offenders who have completely desisted by the beginning of the follow-up period and those who remain active has considerable value.

Brame, Bushway, and Paternoster (2003) illustrated another way to estimate the prevalence of abrupt, permanent, and complete desistance. Their analysis of serious juvenile offenders in the 1958 Philadelphia birth cohort used Poisson and geometric event count models to study the problem of zero-inflation in the distribution of their adult police contacts. Basically, the Brame, Bushway, and Paternoster (2003) analysis was premised on the idea that the event count distribution of active offenders could be explained by a set of discrete rates. Net of these rates, the number of zeros that exceeds those expected by either a mixed Poisson or a mixed geometric distribution provides an estimate of the desistance rate. Although the Poisson and geometric distributions make somewhat different assumptions about the waiting time between contacts, the two specifications yielded similar desistance estimates (Poisson estimate of 36.6 percent and geometric estimate of 38.3 percent) in the Philadelphia data set. It is important to note that the number of people with zero adult contacts was much larger (61.2 percent), which implies that many of the adults with zero contacts were actually still active offenders in adulthood even though they experienced no contacts up to 27 years of age. Clearly, the population of individuals with zero offenses during any finite follow-up period is a superset of those who have completely and permanently desisted at the beginning of the follow-up period. Hence, the split-population survival models considered in this article are directly analogous to the zero-inflated event count models discussed in Brame, Bushway, and Paternoster (2003).

2. Maltz (1984: 113) did note that “when the population under study consists of such poor risks that essentially everyone is eventually expected to fail,” the split-population model would not be reasonable to use. However, this is an empirical question because such samples will yield limiting failure rates that are close to one. When this scenario happens, the split-population model simplifies to the standard waiting-time approach, which imposes the constraint that all cases are “active offenders” at the beginning of the follow-up period and will eventually fail.

An issue that develops in all of the studies discussed thus far (with the notable exception of Kitchener, Schmidt, and Glaser [1977]) is that the follow-up periods are no more than a few years long. This issue may be an artifact of the greater availability of this type of data, or of a system need to assess the immediate risk an offender poses to the public, or both. However, the fact that some people recidivate years after an event and some people do not is an equally compelling, yet understudied, phenomenon. In our view, split-population models combined with long-term follow-up data are an untapped resource that can be deployed to make progress on the question of whether some people truly instantaneously desist, even in samples of serious offenders. In this study, we apply these models to a unique data set of individuals with felony convictions with a long-term follow-up period to understand desistance better as either an instantaneous event or a glide-path toward zero, keeping a wary eye on the possible confounding influence of the changing composition of the sample.

THE ESSEX COUNTY DATA SET

Our analysis is based on the experiences of 962 offenders sentenced as a result of felony convictions or guilty pleas between May 1976 and June 1977 in Essex County, NJ (Gottfredson, 1999a, 1999b). We obtained the data set from the Inter-University Consortium for Political and Social Research (ICPSR #2857) and have analyzed the data as they were archived by the original investigator. Each of these 962 offenders was followed at least until October 1995 when extensive national and state criminal history checks were conducted. These checks continued through February 1997. The dependent variable is the number of days (that we rounded off to the nearest integer number of months) between the sentence date and the first arrest, which is found in the “DAYSTOAR” variable in the archived data set.

Four complications developed as we began our analysis: 1) The data set does not always indicate the point within the May 1976–June 1977 interval when each offender was sentenced; 2) the data set does not disclose when each individual’s record search was conducted; 3) because individuals entered the study at different times in the May 1976–June 1977 period and record searches were conducted at (unknown) different times in the October 1995–February 1997 period, it seems likely that individuals were followed for different lengths of time (ignoring time incarcerated); and 4) the data do not permit us to discern reliably which days individuals were confined in a secure institution such as a youthful offender facility, a jail, or a prison during the follow-up period.

To address these ambiguities, we standardized the follow-up period for all offenders to be 219 months (18.25 years) long, which is the number

of months between June 1977 (the latest enrollment date) and October 1995 (the earliest record search date). All but three individuals who were rearrested had a failure date within the 219-month follow-up interval. For the purposes of our study, the three people who failed after 219 months were treated as censored (meaning that they were not observed to fail within 219 months).

For several reasons, our inability to identify “institutional time” reliably within the 219-month follow-up period does not greatly concern us. First, we note that this problem is not unique to our data set. In fact, most recidivism studies in the field do not include a clear or complete set of confinement data. As an example, consider the North Carolina recidivism study conducted by Schmidt and Witte (1988). This study defined recidivism as a return to a North Carolina state prison at any point during the follow-up period. But offenders can be “off the street” *and* not in a North Carolina state prison in many ways. To the extent that offenders released from prison are arrested on new charges that place them in a North Carolina county jail, those offenders are not “on the street” during that time. To amplify this point, the state’s largest city—Charlotte, NC—has a combined metropolitan statistical area that straddles the North and South Carolina state lines and includes six NC counties and three SC counties (U.S. Office of Management and Budget, 2008: 101). It is plausible that some North Carolina prison releasees in the Charlotte metropolitan area spent some of the follow-up period in county jails in both states and perhaps even spent time in a South Carolina Department of Corrections facility. However, none of this information is measured in Schmidt and Witte’s (1988) data set. In our view, the sorts of issues originating in Schmidt and Witte’s study are actually quite typical of all recidivism studies.

Second, to the extent that confinement suppresses charges for new crimes, it shares that distinction with many other factors not measured in conventional recidivism studies that may have smaller, equal, or more powerful effects on recidivism (e.g., probation and house arrest). If our goal is to understand the waiting-time distribution between sentencing and recidivism, we take it as axiomatic that many factors (most unmeasured) will have important effects on that distribution. Knowing this is the case does not minimize the value of descriptive work documenting the structural features of that distribution.³

3. To address the possibility that time confined significantly distorts our conclusions about the waiting-time distribution between sentencing and rearrest, we estimated all of our statistical models with both the full sample and then with a “reduced sample” of 687 offenders in the Essex County data set who were sentenced to neither a state prison ($n = 182$) or state youth institution ($n = 93$) at the enrollment sentence (the sentence imposed between May 1976 and June 1977, which triggered

Third, our interest in desistance is, by definition, about the right-hand tail of the distribution. With a follow-up period of more than 18 years, we can be assured, even without exact prison data, that the overwhelming majority of our sample had the potential to be free after initial release for a substantial fraction of the observation period. The exact timing of early failure will affect our model choice, but it will not fundamentally affect our analysis of fit during the later periods. To make sure that our modeling exercise is not “misled” by what happens early in the follow-up period, we will pay particular attention to the graphical comparisons of the model in addition to the quantitative measures of model fit used to guide model selection.⁴

Another key feature of our analysis is that we randomly divided the original sample of 962 cases into two groups: 1) an analysis sample with 481 cases and 2) a validation sample with 481 cases. We took this step to reduce the risk of overfitting our model to peculiar or randomly occurring idiosyncratic features of the data set (Schmidt and Witte, 1988). Overfitting is a particular concern in a study like ours where the comparative fit of different statistical models on the same data set is the primary issue. To address this problem, we fit the models on the analysis sample, and then we compare the fitted models with the data from the validation sample. Although not commonly done in criminology, this approach allows us to be more confident about which model best describes the data (Schmidt and Witte, 1988). In particular, it allows us to verify that we can arrive at the same substantive conclusions in both the sample on which the analysis

offenders’ enrollment in the study) (for additional discussion of this distinction in the Essex County sample, see Gottfredson, 1999b: 82). These 687 offenders received a noncustodial ($n = 405$), split ($n = 93$), or jail ($n = 189$) enrollment sentence. Although some differences exist with this reduced sample, the basic conclusions are replicated with the smaller sample—the results are available from the corresponding author. To identify these subsamples, we used the variable INSTSEN in the archived data set. Codes 0 (no sentence to incarceration), 1 (all suspended), and 6 (deferred sentence) were coded as “noncustodial.” Code 2 identifies cases receiving a split sentence in Essex County. Offenders sentenced to jail in Essex County at the enrollment conviction received a code of 3. Youth offenders sentenced to the Yardville facility received a code of 4. Codes 5 (state prison) and 7 (parole violation returned to state prison) were used to identify offenders sent to prison at enrollment.

4. Schmidt and Witte’s (1989) description of their analysis heightens our awareness of this problem. They chose a lognormal model as the best fitting model, but they conceded that the “churn” among people exiting prison/jail together with record-keeping challenges might have accounted for this model choice. Given our focus on differentiating between desistance models, we want to avoid choosing models based on administrative “churn” rather than on individual behavior. The long follow-up, together with our graphical comparisons, lead us to believe that we have avoided this potential pitfall.

is based and a similarly situated sample on which the analysis was not based.

Within the analysis sample, two groups of offenders exist: 1) the 127 (26.4 percent) who received a state prison or state youth institution sentence and 2) the 354 (73.6 percent) who received a noncustodial, jail, or split sentence. Within the validation sample, 148 (30.8 percent) offenders received a state prison or state youth institution sentence, whereas the remaining 333 (69.2 percent) received a noncustodial, jail, or split sentence. The “reduced sample” of 687 cases is the sum of the 354 and 333 cases that received noncustodial, jail, or split sentences at enrollment in the analysis and validation samples, respectively. Figure 2 presents a schematic diagram showing how the Essex County cases fit into these different subgroups.

Next, figure 3 provides a comparison of the hazard rates and cumulative distribution functions for the full sample and the reduced sample. As anticipated, the two samples are similar with the hazard for each rising to a peak at 12 months and then following the well-documented pattern of marginal decline.⁵ For the sake of parsimony, and because we find no major differences between the two samples, we will present the analyses performed on the full sample noting key similarities and differences discovered in the analyses of the smaller subsample.⁶

Figure 4 then provides a similar description of the full sample, broken down into the analysis and validation sections. The plots have a generally similar shape although we note that the hazard rate rises to a higher peak in the validation sample than in the analysis sample. As expected from prior research, the hazard rates in both the analysis and validation samples exhibit an early peak followed by a long decline and then a relatively stable hazard rate near zero. This plot does not seem to have any structure, so we conclude that our random assignment to the analysis and validation samples has produced statistically similar groups.

The right side of figure 4 presents the cumulative distribution functions for the analysis and validation samples. This statistic measures the proportion of the total number of cases ($n = 481$ in the case of each sample) rearrested at or before each time point (up through 218 months as cases are censored at 219 months). The two samples yielded very similar cumulative distribution functions (analysis sample cumulative failure rate = .746; validation sample cumulative failure rate = .761), again suggesting that

5. These hazard rates were calculated by dividing the number of people who were rearrested during each of the 219 months of follow-up by the total number of people who had not yet failed at each time point.

6. All analyses for the subsample were performed analogous to those for the full sample and are available from the authors upon request.

Figure 2. Schematic Description of Full, Reduced, Analysis, and Validation Samples in the Essex County, NJ Data Set ($n = 962$)

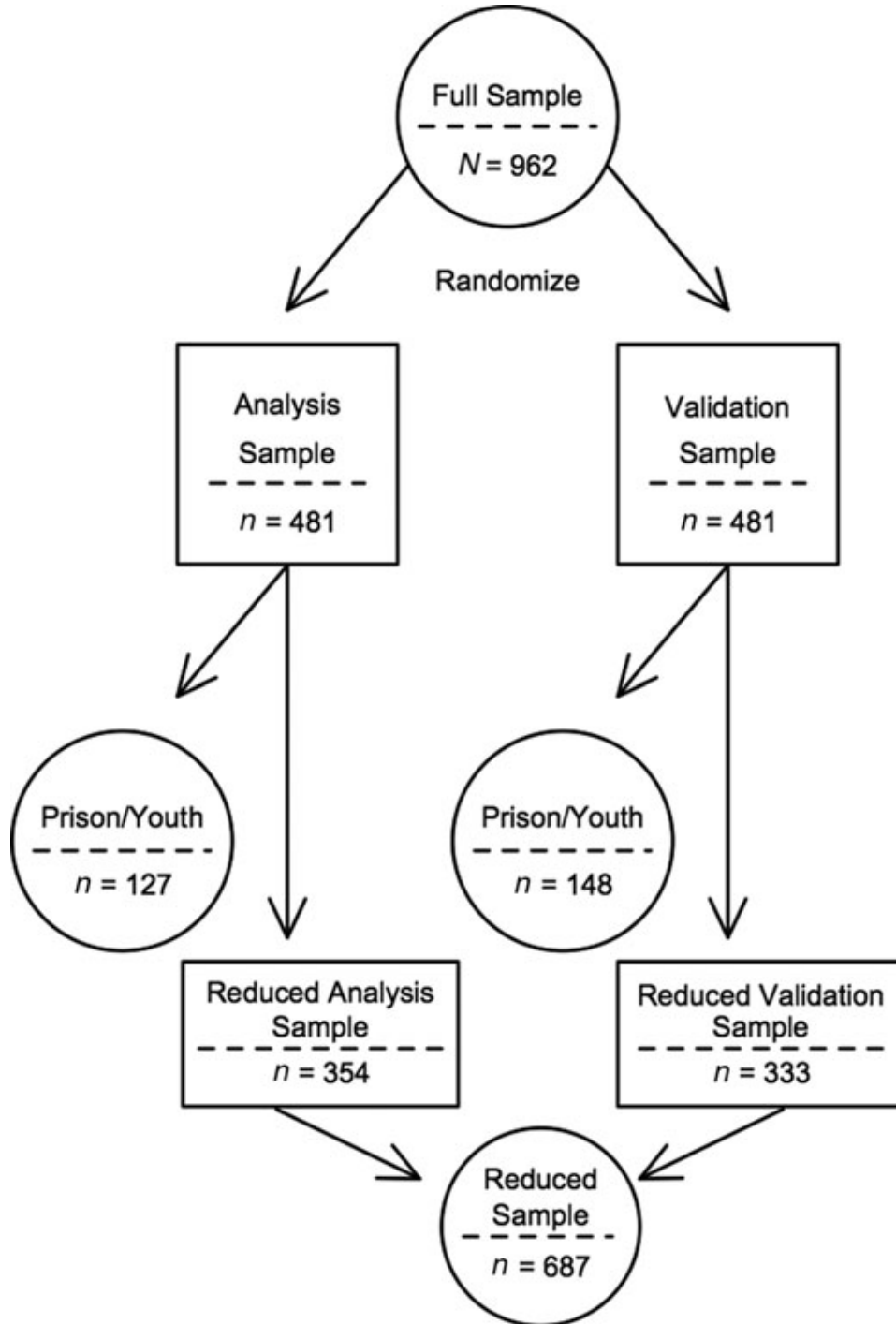


Figure 3. Hazard Rate and Cumulative Proportion Rearrested for Full Essex County Sample ($N = 962$) and Reduced Sample ($n = 687$)

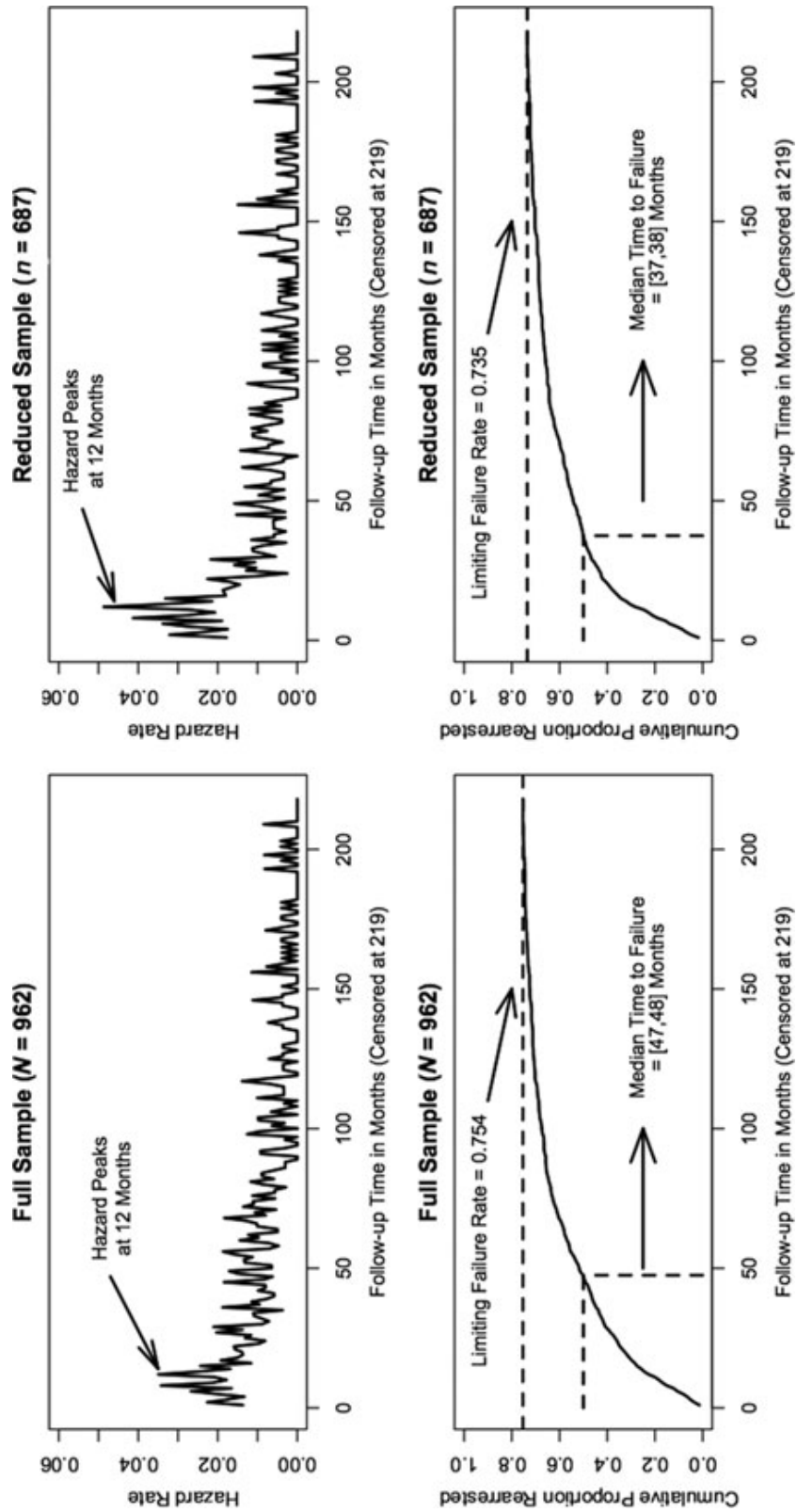
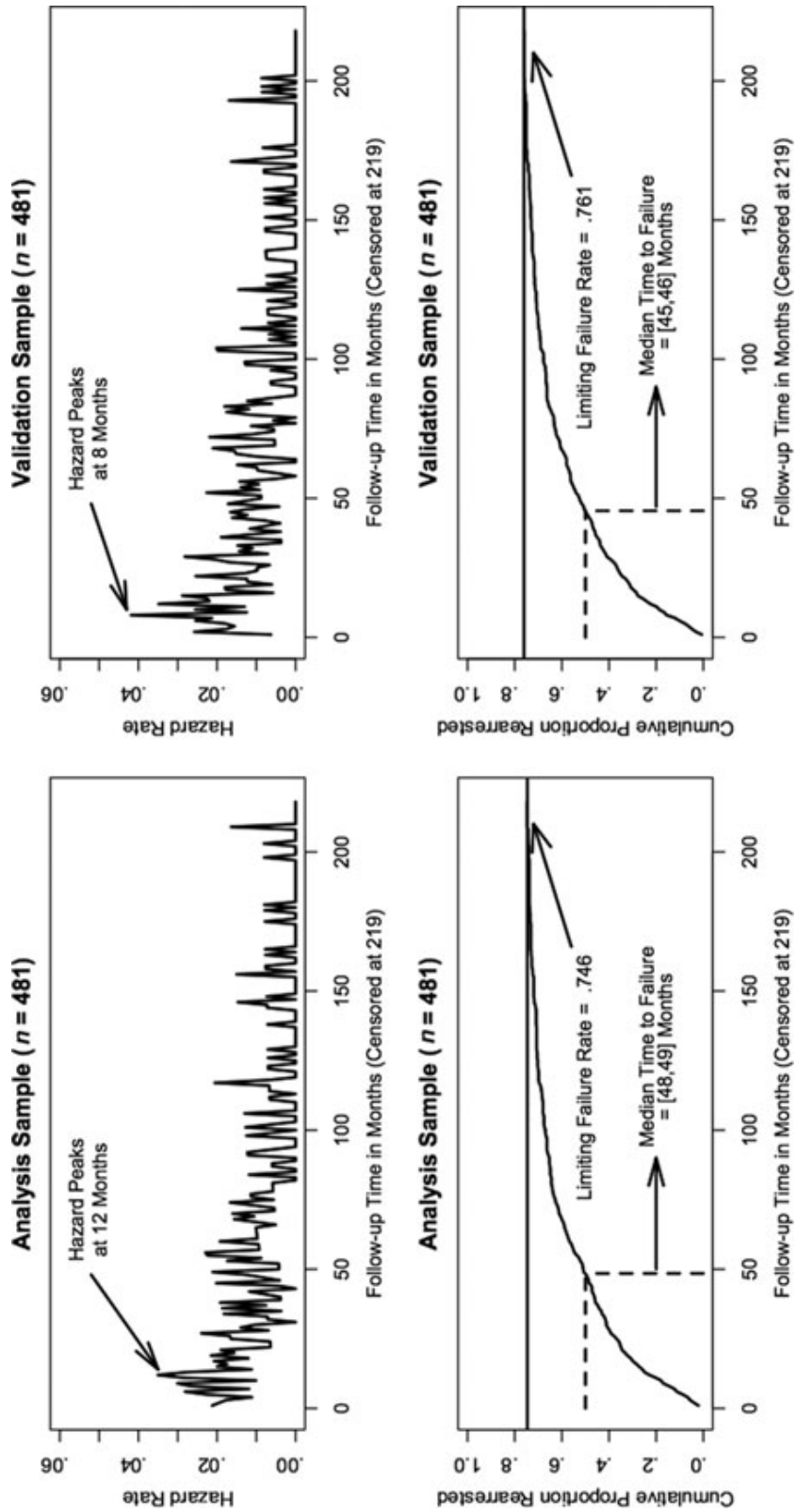


Figure 4. Hazard Rate and Cumulative Proportion Rearrested for Full Essex Sample ($N = 962$) Randomly Divided into Analysis ($n = 481$) and Validation ($n = 481$) Data Sets



the two samples are reasonably comparable with each other. Although not presented in graphical format here, the analysis of the reduced sample produced similar results between its analysis and the validation sample.⁷

ANALYTIC STRATEGY

Our strategy is to estimate several parametric survival-time models to observe what they can tell us about the structure of the Essex County recidivism patterns. The simplest parametric model assumes that the waiting time between sentencing and recidivism follows an exponential distribution, which assumes a time-constant hazard rate. Based on our initial review of the hazard rates and cumulative recidivism distributions in Essex County, it is clear this model does not perform well.

The split-population modification of the exponential model along the lines first suggested by Maltz and McCleary (1977) and discussed further by Maltz (1984) and Schmidt and Witte (1988) is a more plausible specification. This model makes two key assumptions about recidivism and desistance: 1) A group of offenders has already permanently desisted at the time of the enrollment sentence and 2) for the other offenders who remain active at the time of the enrollment sentence, the recidivism hazard rate is time-constant. A noteworthy feature of this model is its incompatibility with the idea that the active offenders' risk of new offending changes over time. The split-population exponential model comprises two parameters: 1) the proportion of offenders who remain active at the beginning of the follow-up period (and the complement of this proportion represents the fraction of the population that has desisted by the beginning of the follow-up period) and 2) the time-constant hazard rate for those who are active offenders at the beginning of the follow-up period.

Based on the recidivism literature, we find four other parametric models as viable alternatives to the split-population exponential model. Two of these models—the Weibull and lognormal—also have two parameters. The Weibull model allows for either a monotonically increasing or monotonically decreasing hazard rate. The lognormal model relaxes the assumption of monotonicity and allows for a turning point in the hazard. Based on the distribution observed in our data, we expect to observe an initial increase in the recidivism hazard followed shortly thereafter by a peak or turning point in the hazard rate. After that point, the hazard rate should drop and flatten with a long right-hand tail. It is important to note that both

7. The only notable exception is that the hazard rate for the nonprison/nonyouth institution analysis group exhibits a sharper peak than what was observed in the full sample analysis. This result is likely because in the full sample, a portion of the individuals were still in physical custody during this early period.

of these models allow for a marginally declining risk of recidivism, but they provide little further information regarding whether this effect is a result of changes in the composition of the sample (instantaneous desisters) or of a gradual decline in the individual risk of recidivism (glide-path to desistance).

The other two models—each with three parameters—are split-population versions of the Weibull and lognormal specifications. Like the split-population exponential model, they both assume some offenders have desisted by the time of the enrollment sentence. Thus, they allow for instantaneous desistance. Unlike the split-population exponential model, however, they relax the constraint of a time-constant hazard rate for the offenders who remain active. Thus, they are compatible with the notion of glide-paths toward desistance. The interpretation of this pattern is still confounded by the changing nature of the sample. More specifically, although the splitting parameter allows for instantaneous desistance and thus accounts for this special case of the compositional effect, the sample is still changing because it is reduced at each time period by those people who failed in the last time period. The declining pattern could be the result of individual change (glide-paths toward desistance) or an artifact of the higher risk offenders “failing” and leaving only lower risk offenders behind. Among the split-population models, only the exponential model provides a strong basis for identifying the different desistance patterns.

The parametric survival-time models used in our study require us to maximize a well-defined log-likelihood function. Our analyses rely on the `maxLik` library that is available in R (version 2.13; Henningsen and Toomet, 2010). The R code and the data used for this analysis are presented in the Appendix.⁸ For both the full sample ($N = 962$) and the reduced sample ($n = 687$), we estimated the above-described six parametric survival-time models of the waiting-time distribution in the analysis subsamples (for the analysis and validation samples in the full sample, $n = 481$).

After estimating these models, we consider several quantitative benchmarks that can be used as the basis for model selection. The first benchmark is the sum of the natural logarithm of the likelihood function evaluated at the maximum likelihood solution for each model. This benchmark is presented for both the analysis and validation samples. The likelihood is an intuitive benchmark for model selection because it is proportional to the joint probability of the waiting-time data looking the way it does conditional on the maximum likelihood solution. The higher the log-likelihood, the more probable the data are conditional on the parameter estimates. The

8. This additional supporting information can be found in the listing for this article in the Wiley Online Library at <http://onlinelibrary.wiley.com/doi/10.1111.crim2012.50.issue-1/issuetoc>.

main drawback of using the log-likelihood alone is that more complex models usually yield higher log-likelihood values. In the case of two models—one that is simpler and one that is more complicated—we need some way to determine whether the improvement offered by the more complex model is large enough to justify rejecting the simpler model. This decision can be difficult if the simpler model performs nearly as well as the complex model.

Two other approaches try to address this problem through the use of penalized likelihood functions. The first approach—the Akaike Information Criterion (AIC)—is obtained by subtracting the number of parameters estimated in a particular model from the sum of the log-likelihood function for that model. Thus, AIC penalizes the log-likelihood for the number of parameters estimated. According to Wasserman (2000: 102), AIC selects the model that produces a density estimate that is close to the density used to generate the data set. The larger the AIC benchmark, the closer the density estimate will be.

In the second approach, the Bayesian Information Criterion (BIC) for a particular model is obtained by dividing the number of model parameters by 2, multiplying the quotient by the natural logarithm of the sample size, and subtracting the product from the log-likelihood function for that model.⁹ Like AIC, the BIC penalizes the log-likelihood for the number of parameters estimated. The BIC, however, also penalizes the log-likelihood for the size of the sample. Hence, as sample sizes grow large, BIC tends to gravitate toward simpler models. With a noninformative prior distribution (i.e., an expression of ignorance about which models in the model space are more probable in an unconditional sense), the BIC (in approximation) will be larger for models that are more probable after conditioning on the data (Wasserman, 2000: 99–102).

In addition to quantitative model selection benchmarks, we also consider graphical evidence of fit. Specifically, we look closely at hazard rate and cumulative distribution function plots to discern the circumstances in which the various models perform well and where they break down. As we discussed, our interest in the right-hand side of the distribution suggests that graphical analysis can and should play a prominent role in deciding which model(s) provide the best fit to the Essex County waiting-time distributions.¹⁰

9. Several equivalent formulas can be found in the literature for calculating AIC and BIC. We used the formulas described by Wasserman (2000: 102).

10. In earlier versions of this article, we experimented with period analysis, an ad hoc statistical approach whereby hazard models are estimated only for people who survive to a certain number of years (e.g., $t = 2$). This method, by construction, focuses attention only on the right-hand tail. These analyses provided additional support for the superior fit of the split-population model in the later years of the

Table 1. Model Fit Benchmarks

Statistical Model	Number of Parameters	Estimation Sample (<i>n</i> = 481)			Validation Sample (<i>n</i> = 481)
		Log(L)	AIC	BIC	Log(L)
Exponential	1	-2,065.513	-2,066.513	-2,068.601	-2,090.744
Split-Exponential	2	-1,967.495	-1,969.495	-1,973.671	-2,000.194
Weibull	2	-2,007.611	-2,009.611	-2,013.786	-2,040.964
Split-Weibull	3	-1,966.800	-1,969.800	-1,976.064	-2,000.062
Lognormal	2	-1,975.291	-1,977.291	-1,981.467	-2,005.834
Split-Lognormal	3	-1,961.945	-1,964.945	-1,971.209	-1,991.084

ABBREVIATIONS: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

ANALYSIS RESULTS

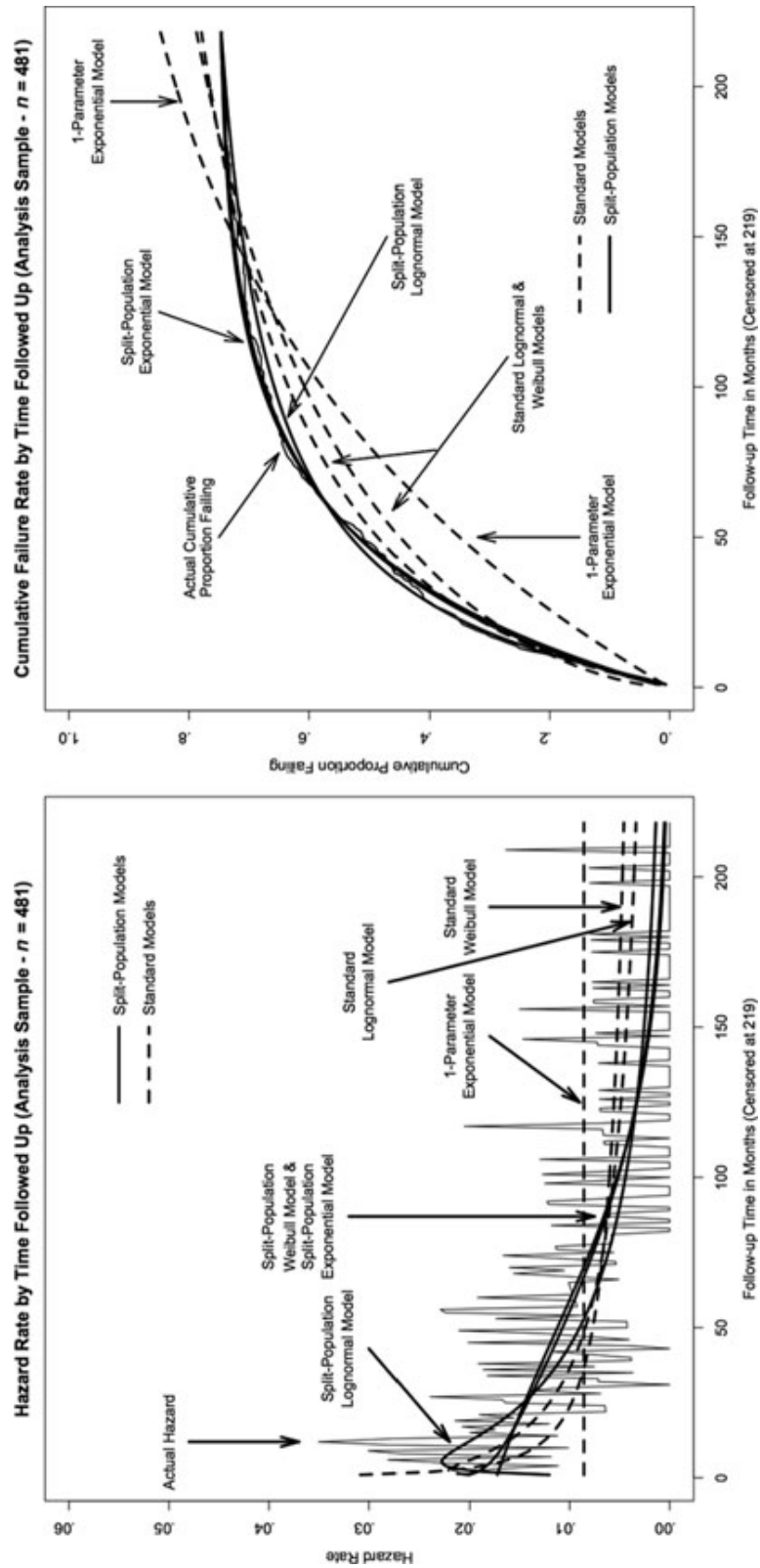
Table 1 presents a summary of the various quantitative model selection benchmarks for each of the six models we estimated in the full samples ($N = 962$).¹¹ The one-parameter exponential model clearly performs the worst, regardless of benchmark. In the full analysis sample, the exponential model yields an estimated constant hazard rate of .0086, which implies a mean waiting time of $1/.0086 = 116.279$ months (Schmidt and Witte, 1988: 53). The median time to failure implied by the exponential model is 80 months compared with the actual medians of [48, 49] months (analysis sample) and [45, 46] months (validation sample). The cumulative failure rate implied by the model at 218 months is 84.8 percent compared with the observed terminal rates of 74.6 percent and 76.1 percent in the analysis and validation samples, respectively.

Overall, as figure 5 shows, the problem with the exponential specification is that the constant hazard rate forces an underprediction of recidivism during the early months of the follow-up period and an overprediction during the later months (see especially the bottom half of the figure for an overview of this deficiency). The time-constant hazard rate is clearly a problem in both the analysis and validation samples. This pattern has been

follow-up period. However, we ultimately concluded that the same insight can be obtained by focusing on the graphical descriptions presented herein without the complications of an additional statistical technique that raises additional questions regarding sample selection bias and truncated likelihood functions.

11. The log-likelihood can be calculated for both the analysis and validation samples using the maximum likelihood solution from the analysis sample. However, as no parameters were estimated for the validation sample, the penalties drop out of the formulas for AIC and BIC leaving only the log-likelihood value. Therefore, we present only the log-likelihood values for the validation sample.

Figure 5. Actual and Predicted Failure Rates in Analysis Sample ($n = 481$)



observed time and again in recidivism studies (Maltz, 1984; Schmidt and Witte, 1988). It seems unlikely that the simple one-parameter exponential model will ever be of much use in modeling criminal recidivism.

The Weibull model performs better than the exponential model. And, as expected, it identifies a monotonically declining hazard rate, which leads to a better estimate of the median time to failure ([62, 63] months) but one that still implies a chronic underprediction of recidivism during the early part of the follow-up period. By the end of the follow-up period, the estimated cumulative failure rate of 78.8 percent is closer than the exponential to the observed value of 74.6 percent in the analysis sample and of 76.1 percent in the validation sample. Despite these improvements, however, it is noteworthy that the Weibull model fits the data much worse than the other two-parameter models (the lognormal and the split-population exponential specifications). Consequently, we do not give any further consideration to the Weibull model either.

Considering the quantitative benchmarks, very little basis exists for distinguishing between the other models. The two-parameter lognormal model and the split-population models all perform well in both the analysis and validation samples.¹² And these models all perform much better than the exponential and Weibull specifications. We pay particular attention to two of these models: 1) the three-parameter split-population lognormal model (because it performs better on all the quantitative benchmarks than the other models) and 2) the two-parameter split-population exponential model (because it performs almost as well as the split-population lognormal model while providing a highly simplified framework for thinking about recidivism and desistance). Figures 6 and 7 provide graphical representations of the cumulative failure residuals based on these models. The one-parameter exponential model also is plotted on these figures as a point of reference for the significance of the improvement obtained by allowing for the splitting parameter. These residual plots are useful for revealing where in the follow-up period the respective models perform best and worst.

The split-population lognormal model projects a limiting (as follow-up times move toward infinity) recidivism rate of 80.1 percent that, in turn, implies that 19.9 percent of the offenders had permanently desisted at the time of the enrollment sentence in 1976 or 1977. Unlike the exponential and

12. A likelihood-ratio test reveals that the three-parameter split-population lognormal model is a significant improvement over the two-parameter lognormal distribution in the analysis sample. The likelihood ratio test is obtained by calculating the difference between the two log-likelihoods and by multiplying that difference by 2. The test statistic is compared with a chi-square distribution with 1 degree of freedom. As no parameters are estimated in the validation sample, the likelihood ratio test is not defined for that sample.

Figure 6. Cumulative Failure Proportion Residuals with Analysis Sample ($n = 481$)

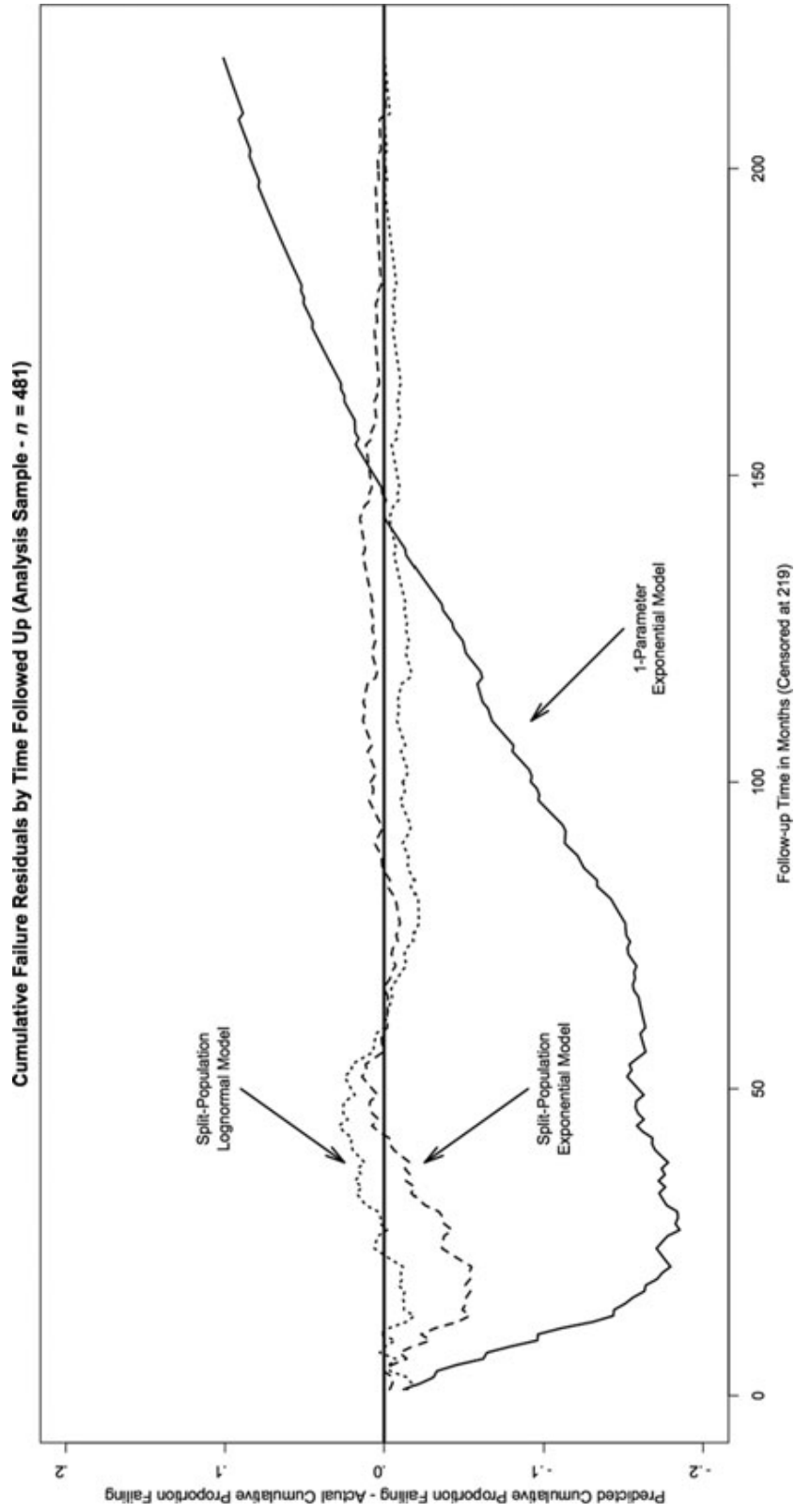
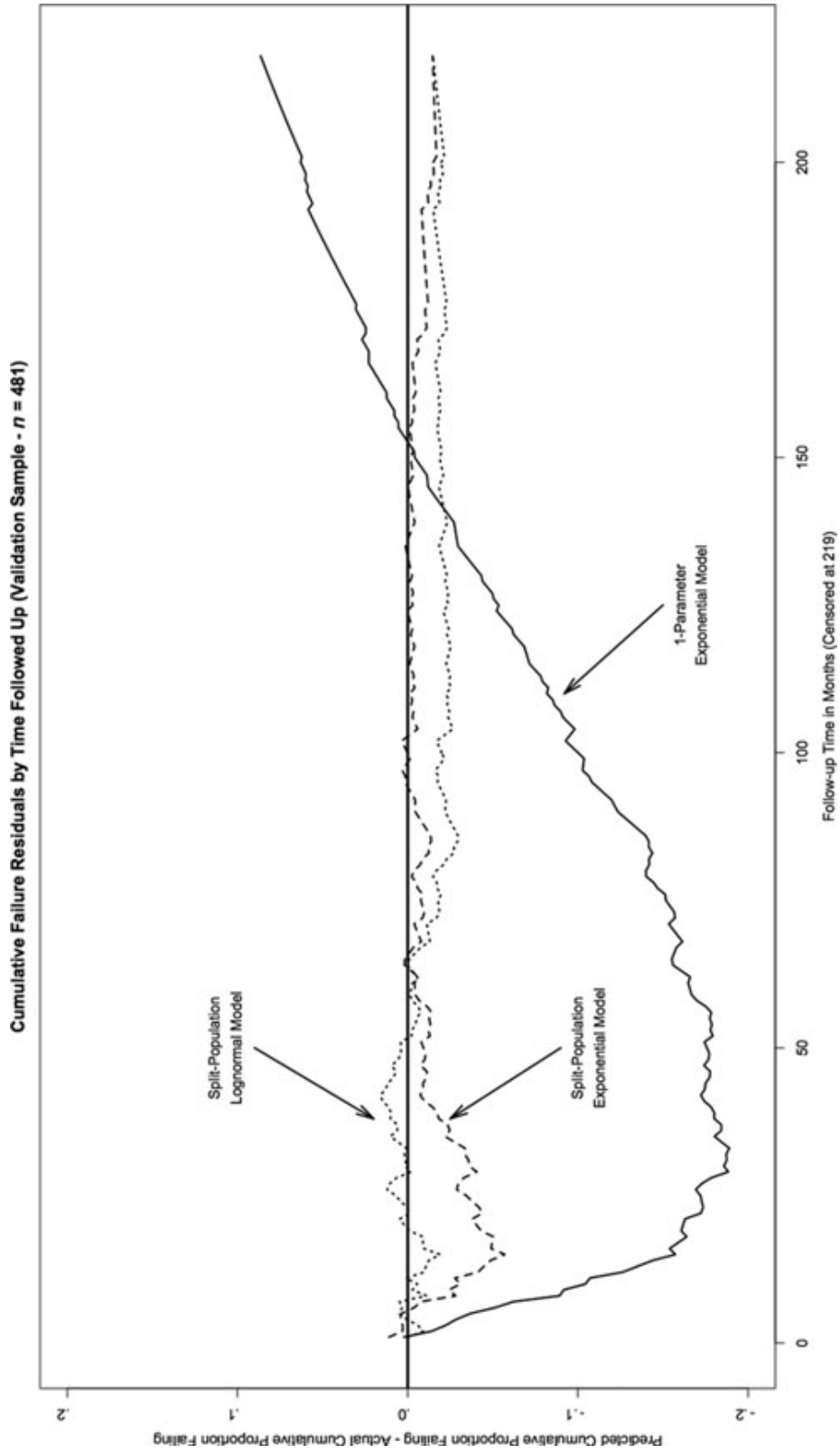


Figure 7. Cumulative Failure Proportion Residuals with Validation Sample ($n = 481$)



Weibull models, which overestimated the median time to failure, the split-population lognormal model yields a median estimate that is too low ([43, 44] months vs. [48, 49] months in the analysis sample and [45, 46] months in the validation sample). The estimated failure rate at 218 months (74.6 percent) is an exact match to the failure rate observed in the analysis sample. A key feature of the split-population lognormal model is its identification of a turning point in the hazard rate early in the follow-up period (the estimated hazard rate peaks at the 5- to 6-month mark at .023). By comparison, the hazard rate peaks in the analysis sample at the 12-month mark (.035), and in the validation sample, the peak occurs at the 8-month mark (.042), which are both later than what the model predicts.

The split-population exponential model predicts a limiting failure rate of 75.1 percent, which is very close to the cumulative recidivism rate at the end of the follow-up period. It follows that if the split-population exponential model is correct, virtually everyone who was going to recidivate has done so by the end of the follow-up period.¹³ Symmetrically, then, the split-population model estimates that a larger percentage of offenders had desisted (24.9 percent) by the time they were initially sentenced in 1976 or 1977. The estimated hazard rate for the active offenders is estimated at .0233, which implies a mean waiting time to recidivism of $1/.0233 = 42.918$ months. This estimate is much lower than what we obtained from the one-parameter exponential model.¹⁴ The split-population exponential model actually does a slightly better job (than the split-population lognormal model) of estimating the median time to failure ([46, 47] months) compared with the actual values of [48, 49] months and [45, 46] months in the analysis and validation samples, respectively.

After comparing the results in figures 6 and 7, it seems reasonable to conclude that the split-population lognormal model fits the data slightly better over the first 3–4 years of follow-up, whereas the split-population exponential fits slightly better after that point. This finding is in and of itself telling. Thinking back to the nature of the models, the lognormal is the only distribution that allows for a turning point. Therefore, the fact that it is fitting the data better in the beginning of the follow-up period signifies that it is gaining much of its “power” in the ability to fit the initial upswing,

13. By way of comparison, if the split-population lognormal model is correct, then some offenders who will ultimately fail have not done so by the end of the follow-up period.

14. This lower result is because the estimated hazard rate for the one-parameter exponential model applies to the whole population, whereas the estimate from the split-population exponential model applies only to the subset of offenders who remained active at the time of the enrollment sentence (Schmidt and Witte, 1988: 70).

peak, and turning point better in the distribution. The fact that the split-exponential model then fits better later in the distribution (in the right tail) is suggestive of an instantaneous model of desistance rather than of a glide-path. Remember that the exponential requires a constant hazard rate; thus, the decline in the hazard would be more indicative of people “failing” and leaving the sample than of individual change in risk of recidivism. If the lognormal, which is a more dynamic model, gained power from its flexibility in the tail of the distribution, this would be indicative of something more than a mere compositional effect—that is, people within the sample are changing. However, the graph shows clearly that, as predicted by Schmidt and Witte (1988), the flexibility of the lognormal is being used for the early years, not for the later years. Simply no strong evidence exists that desisters are taking a long time to get to very low (zero) levels of risk.

As noted, all of the above analyses were performed again on the subsample, which excludes those in prison or youth custody. For this sample, the estimated limiting failure rate is 77.0 percent, which is slightly lower than what was estimated in the full sample (80.1 percent). The median failure time estimated by the model is [37, 38] months, which is shorter than in the full sample. Also, the estimated median is very close to the medians actually observed in the analysis sample (also [37, 38] months) and the validation sample ([36, 37] months). These earlier failure times are most likely because the reduced sample spent less time institutionalized.

These parallel analyses revealed very similar patterns in model fit. The exponential model performs very poorly, whereas the Weibull model performs somewhat better. The split-population lognormal model yields the highest benchmark values for the log-likelihood, AIC, and BIC. Little difference exists on any of the benchmarks among the lognormal, split-population exponential, and split-population Weibull models.¹⁵ For the same reasons as in the full sample, we found the split-population lognormal and split-population exponential models to provide the best fit to the data. However, the split-population exponential model did somewhat more poorly in this analysis than it did in the full sample. The estimated median failure time was 44 months, which is too long. Clearly, the model also is defective between the 1- and 2-year mark where it underpredicts the recidivism rate by approximately 10 percentage points. After that point, the predictions improve, and from the 4-year mark forward, there is little to distinguish between the predictions of the split-population lognormal and those of the exponential specifications. This result is perhaps not surprising given the before noted pattern in the analysis portion of the subsample,

15. As in the full sample, a likelihood ratio test comparing the three-parameter split-population lognormal model with the two-parameter lognormal model indicates that the split-population model is more consistent with the data.

which revealed a sharper peak. If the superiority of the lognormal is indeed in its ability to map this peak and turning point, then it should outperform the split-exponential model to a greater extent when fit to data exhibiting a sharper peak. Both models perform well over the long run, but the three-parameter split-population lognormal model is superior during the early months of the study (particularly between the 1- and 2-year marks).

DISCUSSION AND CONCLUSIONS

Long-term recidivism studies are surprisingly rare in criminology. Given contemporary concerns about crime desistance (National Research Council, 2008) and redemption (Blumstein and Nakamura, 2009; Bushway, Nieuwebeerta, and Blokland, 2011; Kurlychek, Brame, and Bushway, 2006, 2007), we suspect both criminologists and policy makers will begin to place a higher priority on the collection and analysis of long-term recidivism data. Our analysis shows the potential value of this kind of work for testing hypotheses about the processes that lead to desistance.

Our starting point is the widespread understanding—based on decades of recidivism studies—that the risk of offending tends to decline with the passage of time since the last offense. Although the pattern itself has been well understood for decades, what is less clear is why recidivism data usually exhibits this pattern. Two broad sets of explanations exist: 1) that people instantaneous desist from crime at some point, perhaps, having “had enough” or deciding to “go straight” and/or 2) that individuals gradually slow their rate of offending over time as they “glide” toward a point of zero offending. This latter pattern has been theoretically linked to changes in the life course from basic maturation, to the formulation of adult social bonds, to experiential or state-dependent effects. That is, the longer one survives without recidivating, the better one gets at avoiding future criminal behavior.

By building on a mature and well-developed methodological toolbox (Maltz, 1984, 1996; Schmidt and Witte, 1988), we estimated several parametric survival-time models on long-term (18-year) follow-up data from the Essex County, NJ convicted felons study (Gottfredson, 1999a, 1999b). Based on our own findings as well as on previous efforts to model recidivism, we have gained several insights. The first insight is that survival-time models in criminology will most likely always need a parameter that splits the population into two groups: 1) people who have already desisted at the beginning of the follow-up period and 2) those who remain active at the beginning of the follow-up period. These “split-population” models stand in marked contrast to traditional survival-time models, which assume that everyone will fail if followed long enough. Moreover, they imply that at least some of the decline in the risk of recidivism over time is a result

of a particular type of composition. Although criminal career research (Blumstein, Farrington, and Moitra, 1985) has long touted the concept of instantaneous desistance, our study is the first of its kind to note clearly the existence of this group of individuals in a long-term follow-up study. After estimating each of the six parametric models on both the full and subsamples and the analysis and validation samples within each grouping with a follow-up period three times longer than the previous models by Schmidt and Witte (1988), we confidently conclude that instantaneous desistance does exist. Clearly, all split-population models outperformed the simpler models regardless of the specific parametric distribution fitted.

The second insight has two parts. First, the most complicated model—the three-parameter split-population lognormal specification—fits the data best. Second, the much simpler two-parameter split-population exponential specification fit the data almost as well, and indeed, it fit the data better in the later years of the follow-up period. On the surface, this insight may seem to be just a statistical finding, but we argue that each of these models has theoretically relevant content. Because each parametric distribution is based on a key assumption, it is these assumptions that provide substance and meaning to the findings. As pointed out by Maltz (1996), the importance of the underlying model assumptions often is overlooked in the criminological literature. To understand the meaning of our analysis, we must revisit these assumptions.

The split-population lognormal model assumes that two groups of offenders exist: 1) people who have desisted at the beginning of the follow-up period and 2) people who remain active at the beginning of the follow-up period. Within the second group, the split-population lognormal model assumes that the risk of recidivism rises rapidly at the beginning of the follow-up period, reaches a peak, and then declines for the remainder of the follow-up period. As figure 1 of this article and other studies such as those conducted by Schmidt and Witte (1988), Visher, Lattimore, and Linster (1991), and Lattimore and Baker (1993) suggest, a turning point in the hazard rate is not unique to the Essex County data set, but it is a frequently occurring feature of criminal recidivism studies. The fact then that we find the lognormal model to best fit the data is not surprising. However, two more conclusions are gleaned from this study. First, the split-parameter models as previously noted always fit the data better. Thus, a group of instantaneous desisters exists. Second, after accounting for this group, the lognormal model still fits the data better. However, as shown in figure 6, this superiority of the lognormal is confined to the beginning years of the follow-up period.

This finding leads to our next and perhaps most interesting conclusion. That is, the two-parameter split-population exponential model fits the data almost as well as the more complex three-parameter lognormal counterpart

and, in fact, outperforms this model in the later years of the data. It is striking how well this simple model can explain the observed behavior.¹⁶ Like the split-population lognormal model, the split-population exponential model assumes that two groups of offenders exist—those who have desisted at the beginning of the follow-up period and those who remain active. Again, then we find support for instantaneous desistance with the split-population lognormal and exponential model actually reaching similar conclusions about the size of the permanent desisting population at the outset of the follow-up period (the lognormal model is in the 20–23-percent range, whereas the exponential is in the 25–27-percent range). This estimate is smaller than the estimates from Brame, Bushway, and Paternoster (2003) looking at desistance after an arrest. However, it is still substantial. Although the focus of most recidivism studies is on the high recidivism rates, the flip side here is that a full quarter of the sample of felony offenders desists after this conviction. Clearly, not all individuals are equally risky after conviction. Indeed, because the exponential model assumes that the active offenders experience a constant risk of recidivism throughout the follow-up period, no evidence is available of declining hazard rates among the active offenders.

The most profound implication of this model is its suggestion that a decline in the risk of recidivism over time is an artifact of different groups that behave in highly stable ways over the course of the follow-up period. Therefore, the fact that the split-population exponential models fits the data better in the later years provides greater support for the traditional criminal career notion of stable rates of offending over time combined with a fixed probability of desistance in each period rather than our proposed alternate explanation of a glide-path toward desistance as individuals gradually slow in their rate of offending.

It seems evident that the length of the follow-up period in the Essex County data set has a lot to do with our assessment of how well the split-population exponential model performs. If the Essex County study had only followed offenders for 3, 4, or 5 years—typical follow-up periods for recidivism studies—our conclusions about the split-population lognormal and exponential models would have been different. Over this window of time, the split-population lognormal model clearly performs better. But viewed over the entire 18-year follow-up period, the simpler, two-parameter split-population exponential model emerges as a formidable competitor. As

16. The result is all the more striking given that these are unconditional models with no observable variables included to control for population heterogeneity (sorting). Not surprisingly, the results hold with control variables. Moreover, in results not reported here, we found no evidence that our results were driven by different parts of the sample as described by risk.

more data sets with long follow-up periods are studied, it will be interesting to observe how well the split-population exponential model performs, especially after the first few years of follow-up.

A final insight revolves around the concept of intermittency or reactivation of criminal careers after a period of dormancy or “temporary desistance” (Barnett, Blumstein, and Farrington, 1989; Horney, Osgood, and Marshall, 1995; Nagin and Land, 1993). The concept of intermittency has been gaining ground in criminology in recent years and leads to certain theoretical and policy implications (for example, the idea that desistance is always provisional). Our analysis is certainly consistent with the idea that a low-rate offender can go for many years before committing a new offense. But intermittency is a particularly dynamic model of offending in which the offender goes from an active rate of offending to a zero rate of offending back to a fully active criminal career. Tests of this dramatic assumption are not supported by the current analysis or data structure.

Barnett, Blumstein, and Farrington (1989) moved to an intermittency explanation after they found evidence of a “fat” tail—higher rates of offending more than 5 years after the last offenses than could be explained by the exponential model. Although we found support for their simple split-population exponential model, we did not find a fat tail even though we had a more serious population and longer follow-up periods. The exponential model can easily accommodate the amount of offending we found later in the observation period. As a result, we conclude that no evidence is available for intermittency, at least as described by Barnett, Blumstein, and Farrington (1989). As data sets with longer follow-up periods become more widely available, we suggest that others also investigate the value of intermittency as a concept.

REFERENCES

- Anscombe, Francis John. 1961. Estimating a mixed-exponential response law. *Journal of the American Statistical Association* 56:493–502.
- Barnett, Arnold, Alfred Blumstein, and David P. Farrington. 1987. Probabilistic models of youthful criminal careers. *Criminology* 25: 38–108.
- Barnett, Arnold, Alfred Blumstein, and David P. Farrington. 1989. A prospective test of a criminal career model. *Criminology* 27:373–85.
- Baskin, Deborah, and Ira Sommers. 1998. *Casualties of Community Disorder: Women’s Careers in Violent Crime*. Boulder, CO: Westview Press.

- Beck, Allen J., and Bernard E. Shipley. 1989. *Recidivism of Prisoners Released in 1983*. Bureau of Justice Statistics Special Report. Washington, DC: U.S. Department of Justice.
- Blumstein, Alfred, Jacqueline Cohen, Jeffrey A. Roth, and Christy A. Visher, eds. 1986. *Criminal Careers and Career Criminals*, vols. 1 and 2. Washington, DC: National Academy Press.
- Blumstein, Alfred, David P. Farrington, and Soumyo Moitra. 1985. Delinquency careers: Innocents, desisters, and persisters. *Crime and Justice* 6:187–219.
- Blumstein, Alfred, and Richard Larson. 1969. Models of a total criminal justice system. *Operations Research* 17:199–232.
- Blumstein, Alfred, and Soumyo Moitra. 1980. The identification of “career criminals” from “chronic offenders” in a cohort. *Law and Policy Quarterly* 2:321–34.
- Blumstein, Alfred, and Kiminori Nakamura. 2009. Redemption in the presence of widespread criminal background checks. *Criminology* 47:327–59.
- Brame, Robert, Shawn D. Bushway, and Raymond Paternoster. 2003. Examining the prevalence of criminal desistance. *Criminology* 41: 423–48.
- Bushway, Shawn D., Paul Nieuwbeerta, and Arjan Blokland. 2011. The predictive value of criminal background checks: Do age and criminal history affect time to redemption? *Criminology* 49:27–60.
- Bushway, Shawn D., Alex R. Piquero, Lisa M. Broidy, Elizabeth Cauffman, and Paul Mazerolle. 2001. An empirical framework for studying desistance as a process. *Criminology* 39:491–513.
- Carr-Hill, G. A. and R. A. Carr-Hill. 1972. Reconviction as a process. *British Journal of Criminology* 12:35–43.
- Dean, Charles W., Robert Brame, and Alex R. Piquero. 1996. Criminal propensities, discrete groups of offenders, and persistence in crime. *Criminology* 34:547–74.
- DeJong, Christina. 1997. Survival analysis and specific deterrence: Integrating theoretical and empirical models of recidivism. *Criminology* 35:561–75.

LONG-TERM CRIME DESISTANCE AND RECIDIVISM 101

- Gottfredson, Don M. 1999a. *Effects of Judges' Sentencing Decisions on Criminal Careers*. NIJ Research in Brief. Washington, DC: U.S. Department of Justice.
- Gottfredson, Don M. 1999b. *Choosing Punishments: Crime Control Effects of Sentences*. Unpublished final report to the National Institute of Justice, Document #179278.
- Henningsen, Arne, and Ott Toomet. 2010. maxLik: A package for maximum likelihood estimation in R. *Computational Statistics* 26:443–58.
- Horney, Julie, D. Wayne Osgood, and Ineke H. Marshall. 1995. Criminal careers in the short term: Intra-individual variability in crime and its relation to local life circumstances. *American Sociological Review* 60:655–73.
- Kitchener, Howard, Annesley K. Schmidt, and Daniel Glaser. 1977. How persistent is post-prison success? *Federal Probation* 41:9–15.
- Klugman, Stuart A., Harry H. Panjer, and Gordon E. Wilmot. 2008. *Loss Models: From Data to Decisions*, 3rd ed. New York: Wiley.
- Kurlychek, Megan C., Robert Brame, and Shawn D. Bushway. 2006. Scarlet letters and recidivism: Does an old criminal record predict future offending. *Criminology & Public Policy* 5:483–504.
- Kurlychek, Megan C., Robert Brame, and Shawn D. Bushway. 2007. Enduring risk? Old criminal records and predictions of future criminal involvement. *Crime & Delinquency* 53:64–83.
- Langan, Patrick A., and David J. Levin. 2002. *Recidivism of Prisoners Released in 1994*. Bureau of Justice Statistics Special Report. Washington, DC: U.S. Department of Justice.
- Lattimore, Pamela K., and Joanna R. Baker. 1993. The impact of recidivism and capacity on prison populations. *Journal of Quantitative Criminology* 8:189–215.
- Laub, John H., and Robert J. Sampson. 2001. Understanding desistance from crime. In *Crime and Justice: A Review of Research*, vol. 28, eds. Michael H. Tonry and Paul Mazzerolle. Belmont, CA: Wadsworth.
- Leblanc, Marc, and Rolf Loeber. 1998. Developmental criminology updated. In *Crime and Justice: A Review of Research*, vol. 23, ed. Michael H. Tonry. Chicago, IL: University of Chicago Press.
- Maltz, Michael D. 1984. *Recidivism*. New York: Academic Press.

- Maltz, Michael D. 1996. From Poisson to the present: Applying operations research to problems in criminal justice. *Journal of Quantitative Criminology* 12:3–61.
- Maltz, Michael D., and Richard McCleary. 1977. The mathematics of behavioral change. *Evaluation Quarterly* 1:421–38.
- Maruna, Shadd. 1999. Criminology, desistance, and the psychology of the stranger. In *The Sociology of Crime: Groups, Teams, and Networks*, eds. David Cantor and L. Alison. Aldershot, U.K.: Ashgate.
- Nagin, Daniel S., and Kenneth C. Land. 1993. Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology* 31:327–62.
- National Research Council. 2008. *Parole, Desistance from Crime, and Community Integration*. Washington, DC: National Academy Press.
- Sampson, Robert J., and John H. Laub. 2003. Life-course desisters? Trajectories of crime among delinquent boys followed to age 70. *Criminology* 41:319–39.
- Schmidt, Peter, and Ann Dryden Witte. 1984. *An Economic Analysis of Crime and Justice: Theory, Methods, and Applications*. New York: Academic Press.
- Schmidt, Peter, and Ann Dryden Witte. 1988. *Predicting Recidivism Using Survival Models*. New York: Springer-Verlag.
- Soothill, Kevin, and Brian Francis. 2009. When do ex-offenders become like non-offenders? *Howard Journal of Criminal Justice* 48:373–87.
- Stollmack, Stephen, and Carl M. Harris. 1974. Failure-rate analysis applied to recidivism data. *Operations Research* 22:1193–205.
- U.S. Office of Management and Budget. 2008. *Update of Statistical Area Definitions and Guidance on their Uses*. OMB Bulletin Number 09–01. Washington, DC: U.S. Office of Management and Budget.
- Visher, Christy A., Pamela K. Lattimore, and Richard L. Linster. 1991. Predicting the recidivism of serious youthful offenders using survival models. *Criminology* 29:329–63.
- Wasserman, Larry. 2000. Bayesian model selection and model averaging. *Journal of Mathematical Psychology* 44:92–107.

LONG-TERM CRIME DESISTANCE AND RECIDIVISM 103

Witte, Ann Dryden, and Peter Schmidt. 1977. An analysis of recidivism using the truncated lognormal distribution. *Applied Statistics* 26:302–11.

Megan C. Kurlychek is an associate professor of criminal justice in the School of Criminal Justice at the University at Albany. She received her PhD in crime, law and justice from Pennsylvania State University in 2004. Her research interests center on the etiology of crime and delinquency and on the ways in which system involvement subsequently either serves to perpetuate, decelerate, or halt future criminality.

Shawn D. Bushway is an associate professor of criminal justice, and public administration and policy at the University at Albany (SUNY). He received his PhD in public policy analysis and political economy in 1996 from the Heinz School of Public Policy and Management at Carnegie Mellon University. His current research focuses on the process of desistance, the impact of a criminal history on subsequent outcomes, and the distribution of discretion in the criminal justice sentencing process.

Robert Brame is a professor in the Department of Criminal Justice and Criminology at the University of North Carolina at Charlotte. Currently, his research interests include measurement of various criminal career parameters such as onset, recidivism, specialization, and escalation. He also is working on a project examining law enforcement and court processing responses to misdemeanor domestic violence.

Appendix

- A. Essex County Full Sample ($N = 962$).
- B. Essex County Reduced Sample ($n = 687$).
- C. Likelihood Function for Exponential Model.
- D. Generate Data for Exponential Diagnostic Plots.
- E. Generate Diagnostic Plots for Exponential Model.
- F. Likelihood Function for Weibull Model.
- G. Likelihood Function for Lognormal Model.
- H. Likelihood Function for Split-Population Exponential Model.
- I. Likelihood Function for Split-Population Weibull Model.
- J. Likelihood Function for Split-Population Lognormal Model.