Article

Mathematical Models of Criminal Careers: Deriving and Testing Quantitative Predictions

David P. Farrington[1], John F. MacLeod[1], and Alex R. Piquero[2]

Abstract
Theories in criminology rarely make exact quantitative predictions that can be tested empirically. This article reviews mathematical models of criminal careers, which are simple theories that fit a wide variety of empirical data. It focuses on the work of Blumstein and his colleagues in the 1980s and on the more recent research of MacLeod, Grove, and Farrington. Criminal career data can be fitted by simple assumptions specifying that the frequency of offending and the probability of recidivism are constant over time and that there are two or three categories of offenders who differ in these parameters. These theories also predict future offending. It is useful to build on these simple mathematical models to predict a wider range of criminological results and convert criminology into a more predictive and accurate science.

[1] Institute of Criminology, Cambridge University, Cambridge, UK
[2] Program in Criminology, University of Texas at Dallas, Richardson, TX, USA

Corresponding Author:
David P. Farrington, Cambridge University, Sidgwick Avenue, Cambridge CB3 9DA, UK.
Email: dpf1@cam.ac.uk

Introduction

Unlike in the physical sciences, theories in criminology, as in the other behavioral and social sciences, rarely make exact quantitative predictions that can be tested empirically. For example, Moffitt's (1993) theory specifies that there are two types of offenders: life-course-persistent, who have an early onset and a long criminal career, and adolescence-limited, who have a later onset and a short criminal career, each of whom follow different pathways because of different influencing risk factors. This theory is superior to most other criminological theories in yielding testable hypotheses and in the extent to which it has been empirically tested (Moffitt 2006). Yet, while it provides some general predictions about offending patterns by the two groups, it does not specifically predict such quantities as the exact fraction of any birth cohort who will become life-course-persistents compared with adolescence-limiteds, the frequency of offending of each category of offenders at each age, the distribution of ages of onset and criminal career durations of each category, or other features of the age–crime curve. In a true science, exact quantitative predictions should be made by theories, which permit them to be tested empirically, and then predictions may be supported, refuted, or modified. A number of theories and mathematical models that make exact quantitative predictions have been proposed in psychology (Atkinson, Bower, and Crothers 1965; Laming 1973) and criminology (Greenberg 1979; Piquero and Weisburd 2010).

In this article, we review simple mathematical models that make exact quantitative predictions about key features of criminal careers. The most important of these models of criminal careers have been inspired by the work of Alfred Blumstein. He defined a criminal career as a longitudinal sequence of offenses committed by an individual offender; thus, the study of criminal careers requires longitudinal data on offending, whether obtained from official records or from longitudinal surveys such as the Cambridge Study in Delinquent Development (CSDD), in which over 400 London boys have been followed up in interviews and records from age 8 to 48–56 (Farrington et al. 2006, 2014; Farrington, Piquero, and Jennings 2013; Piquero, Farrington, and Blumstein 2007).

Table 1. Summary of Criminal Career Models.

Author	Main Assumptions about λ and p	Main Career Features Predicted
Blumstein and Moitra (1980)	Innocents, amateurs, and persisters	Recidivism probabilities
Blumstein et al. (1985)	Innocents, desisters, and persisters	Recidivism probabilities
Barnett and Lofaso (1985)	λ is constant	Future arrest rate
Barnett et al. (1987)	Frequents and occasionals	Number of offenses by offenders and time intervals between convictions
Greenberg (1991)	λ has gamma distribution	Number of offenses by individuals
Rowe et al. (1990)	λ has lognormal distribution	Number of offenses by individuals
Nagin and Land (1993)	λ has continuous distribution	Number of offenses by individuals
MacLeod et al. (2012)	Two categories of λ and two categories of p	Number of offenses, reconviction probabilities, and age–crime curve

Note: λ = individual offending frequency; p = probability of recidivism.

After summarizing the work of Blumstein, and the U.S. National Academy of Sciences panel on Criminal Career Research, we describe the development of some simple mathematical models that predict a wide range of criminal career features. We then review competing models that developed into the group-based trajectory modeling technique of Nagin (2005). Finally, we review the work of MacLeod, Grove, and Farrington (2012), who continued the original approach of Blumstein and his colleagues by proposing and testing very simple models that predicted many different criminal career results. As a roadmap for the reader, Table 1 summarizes these criminal career models, each of which makes different assumptions about the rate of offending, the number of groups in the offending population, either assumed a priori or identified by particular modeling approaches, and the main criminal career dimensions and their predictors.

Early Research by Blumstein and Cohen

In the 1970s, Blumstein became very interested in estimating the incapacitative effects of imprisonment, following the work of Shinnar and Shinnar

(1975), and he chaired the U.S. National Academy of Sciences Panel on Deterrence and Incapacitation (Blumstein, Cohen, and Nagin 1978). He became convinced that, in order to address the key policy issues such as incapacitation, it was crucial to advance knowledge about key features of criminal careers. Blumstein and Cohen (1979) published a landmark paper that addressed many of the most crucial issues in criminal career research.

Blumstein and Cohen used λ to indicate the individual offending frequency (the average number of crimes committed per year by active offenders) and μ to measure the individual arrest frequency (the average number of arrests per year of active offenders). They used the symbol λ because they assumed that crimes occurred at random over time (at a constant rate and independently), which naturally led to the deduction that the number of crimes in any time period had a Poisson distribution with mean λ. Other researchers had also proposed that crimes occurred according to a Poisson process (Carr-Hill and Carr-Hill 1972; Shinnar and Shinnar 1975). Blumstein and Cohen defined λ as the true (underlying) frequency of offending and μ as the measured frequency of offending. λ and μ were linked by q, which was the probability of arrest following a crime:

$$\mu = \lambda * q (\text{where } * \text{ indicates multiplication}).$$

After comparing the number of arrests with the number of crimes recorded by the police and the probability of reporting a crime to the police according to an early victim survey, Blumstein and Cohen estimated that q varied from .025 for larceny to .111 for aggravated assault, while λ averaged 10 index crimes and 24 crimes of all kinds per year per active offender. They pointed out that λ and μ were stochastic variables that varied (over time) randomly about a mean value, so that an active offender (a person in a criminal career, between onset and termination, with a nonzero λ or frequency of offending) nevertheless had a certain probability of committing no crimes in a particular year. They concluded that there was little evidence of specialization in types of crimes and that λ and μ stayed tolerably constant at different ages (within different crime types).

Around the same time that Blumstein and his colleagues were developing criminal career models, a major effort to measure λ was completed by the Rand Corporation, in which nearly 2,200 jail and prison inmates were surveyed in three states and asked about their involvement in offending (Greenwood and Abrahamse 1982). They were interested in estimating the incapacitative effect of incarceration, and this required the estimation of λ, q, J (the probability of incarceration given a conviction), and S (the average time served). They developed a seven-point scale that discriminated

between high λ, medium λ, and low λ offenders (burglars and robbers), and estimated the incapacitative effects of extending sentences for high-λ offenders and reducing sentences for low-λ offenders. For example, for California robbers, they calculated that this policy might simultaneously reduce the robbery rate by 15 percent and reduce the number of persons incarcerated for robbery by 5 percent. These conclusions proved to be quite controversial, and a reanalysis of the Rand inmate survey by Visher (1986) suggested that the effects of a selective incapacitation policy might have been overestimated.

The 1986 U.S. National Academy of Sciences Panel

Blumstein chaired the path-breaking National Academy of Sciences Panel on Criminal Career Research, which documented the criminal career paradigm in great detail (Blumstein et al. 1986). The panel defined many criminal career features: not only λ, μ, and q but also T (criminal career duration), a_o (age at onset), a_T (age at the termination of a career), d (participation rate or prevalence of offending per year), C (the aggregate crime rate per year), and δ (the fraction of active offenders who terminate at each age). The panel also emphasized the need to disentangle aggregate crime rates into prevalence and frequency, bearing in mind the following equation:

$$C = \lambda * d$$

For example, an increase in the aggregate crime rate may be caused by an increase in the prevalence of offenders (d) or by an increase in the frequency of offending by active offenders (λ) or by both. In implementing criminal justice policies, it is important to know what is happening, because an increase in the prevalence of offenders would call for primary prevention targeted at the whole community or secondary prevention targeted at high-risk persons, whereas an increase in the frequency of offending would call for tertiary prevention targeted at the most active offenders. Piquero, Farrington, and Blumstein (2003) provided a summary of the extensive research undertaken on criminal career issues. In the course of that research, criminologists have proposed a number of relevant mathematical models, each making fairly precise predictions about the parameters of interest.

Predicting the Growth in Recidivism Probabilities

In an enormously influential longitudinal study of nearly 10,000 boys born in Philadelphia in 1945 and followed up in official records to the 18th

birthday, Wolfgang, Figlio, and Sellin (1972:162) found that the probability of recidivism increased after each successive offense. For example, it was .54 after the first offense, .65 after the second offense, .72–.74 after offenses 3–5, .77–.79 after offenses 6–7, and subsequently appeared to reach an asymptote of about .80–.83. They also found that 6 percent of the birth cohort (18 percent of the offenders), all of whom had committed at least five offenses, accounted for the majority (52 percent) of all crimes. These 6 percent accounted for even higher proportions of serious crimes: 69 percent of all aggravated assaults, 71 percent of homicides, 73 percent of forcible rapes, and 82 percent of robberies. The discovery of these 6 percent, termed ''chronic offenders,'' led to calls for early identification of those prime targets for intervention and the application of appropriate methods of prevention and treatment. Several other longitudinal studies, including the CSDD in the United Kingdom (Farrington and West 1993), also found that about 5–6 percent of a cohort accounted for over half of all crimes of that cohort.

Blumstein and Moitra (1980) pointed out that Wolfgang et al. (1972) identified the chronic offenders retrospectively. Even on the simplest assumption that every offender in the cohort had the same probability p of reoffending after each crime, chance factors alone would result in some of them having more arrests and others fewer. Because of these probabilistic processes, those with the most arrests—defined after the fact as the ''chronics''—would account for a disproportionate fraction of the arrests. This can be illustrated by an example based on throwing a die. If an unbiased die was thrown 30 times and the 5 highest scores were added up, these would account for a disproportionate fraction of the total score obtained in all 30 throws (30 of 105, on average). In this die-throwing example, picking the highest scores means that, purely by chance, 16.7 percent of the throws would account for 28.6 percent of the total score. In light of this, 18 percent of offenders accounting for 52 percent of offenses does not seem quite so remarkable.

Blumstein and Moitra (1980) proposed a simple model to predict the growth in recidivism probabilities. They assumed that the probability of a first offense, p_1, was .35, since 35 percent of the boys were arrested. Similarly, they proposed that $p_2 = .54$ and $p_3 = .65$, the observed figures. However, they then proposed that the probability of committing a subsequent offense, p_k, was always .72 after every offense from the third onward, or in other words did not vary with the serial number of the offense k. They showed that they could fit the data (the number of boys committing each number of offenses) quite well with this model.

Blumstein and Moitra divided the cohort into "innocents" (those with no offenses), "amateurs" (those with one to three offenses), and "persisters" (those with more than three offenses). Because offenses are committed at random (with a constant probability of .72 after the third onward), the expected number of offenses committed after each offense is also constant and does not vary with the serial number of the offense. Simple mathematics shows that:

$$E = p_k/(1 - p_k) \text{ (where } E \text{ is the expected number of offenses),}$$

When $p_k = .72$, $E = 2.57$ (.72/.28).
Blumstein and Moitra (1980:327) concluded:

> Thus, if we imprison all persons who have already been arrested three times, we avert 2.57 future arrests per prisoner. If we restrict imprisonment to the "more chronic" offenders who have already been arrested seven times, we have to deal with fewer prisoners, but we still avert only 2.57 future arrests per prisoner. Thus, more efficient incapacitation cannot be achieved by using a higher value of "chronicness."

Blumstein, Farrington, and Moitra (1985) then improved on this model by further partitioning the persisters. They assumed that $p_4 = p_5 = p_6 = .72$ and that $p_k = .80$ for every offense after the sixth. This model provided a better fit to the observed data. The constant recidivism probability of .80 meant that, for every offense from the sixth onward, the expected number of subsequent offenses was 4 (.80/.20). Blumstein et al. (1985) also showed that a similar model fitted the data from three other longitudinal surveys, including the CSDD. The problem with this model, however, is that most of the parameters (the probabilities) are set empirically and retrospectively. Blumstein et al. (1985) therefore proposed a more parsimonious model by assuming that there were only three categories of boys in the cohort: (1) innocents, with a zero probability of arrest, who comprise a fraction β of the cohort; (2) "desisters," with a low probability of arrest p_a, who comprise a fraction α of the offenders; (3) "persisters," with a high probability of arrest p_r, who comprise a fraction $(1 - α)$ of the offenders. In this model, the persisters and desisters are distinguished by their prospective (a priori) probabilities rather than by their actual numbers of arrests. There are persisters and desisters with each number of arrests, but as the number of arrests increases the mix of arrestees is increasingly dominated by persisters.

Blumstein et al. (1985) showed that this model fitted the Philadelphia data very well, with the following parameters: β (fraction innocent) = .65,

p_r (recidivism probability for persisters) = .80, p_a (recidivism probability for desisters) = .35, and α (fraction of first offenders who are desisters) = .56. They compared the actual and predicted number of boys with each number of arrests, based on the actual sample size of 9,945 boys. For example, there were 1,613 boys with 1 arrest, and the model predicted 1,571; there were 344 boys with 3 arrests, and the model predicted 351; and there were 39 boys with 10 arrests, and the model predicted 41. This is a good example of a simple theory predicting and fitting empirical data on criminal careers.

Blumstein et al. (1985) then applied their mathematical model (of innocents, desisters, and persisters) to the CSDD data. The best fit to the recidivism probabilities in the survey was obtained by assuming that the probability of persisting after each conviction was .87 for persisters and .57 for desisters. The proportion of first offenders who were persisters was .28, while the fraction of the sample who were innocents was .67. Persisters and desisters differed in their a priori probabilities of persisting, not in their a posteriori number of convictions (as chronics did).

Interestingly, the number of empirically predicted chronics among the offenders (37 "high-risk" offenders with four or more of seven childhood risk factors) was similar to the predicted number of persisters (36.7) according to the mathematical model. Remarkably, the individual process of dropping out of crime by the predicted chronics in the empirical data closely matched the aggregate dropout process for persisters predicted by the mathematical model with parameters estimated from aggregate recidivism data analysis. Therefore, the high-risk offenders might be viewed as the identified persisters. This analysis shows the important distinction between prospective empirical predictions (e.g., high-risk offenders), underlying theoretical categories (e.g., persisters), and retrospectively measured outcomes (e.g., chronics).

Predicting Individual Offending Frequency

Barnett and Lofaso (1985) also analyzed the Philadelphia cohort data but aimed to predict the individual offending frequency rather than the number of offenses committed. They assumed that offenses were committed probabilistically (at random) over time, which means that offenders accumulate arrests according to a stationary Poisson process (with a constant mean rate). They found that the best predictor of the future individual offending frequency (crimes per year) was the past individual offending frequency. These authors also drew attention to the problems caused by truncation of the data at the 18th birthday. Recidivism probabilities for juveniles are

often calculated on the assumption that all active offenders desist after their last juvenile arrest. However, especially in the case of a boy who was arrested shortly before his 18th birthday, this "desistance" might be false, since it is quite likely that he would have a subsequent adult arrest. Assuming that arrests occurred probabilistically according to a Poisson process, they calculated the probability of no arrest occurring between the last juvenile arrest and the 18th birthday, given that the offender was continuing his criminal career and had not truly desisted. Taking account of the past arrest rate and the time at risk between the last arrest and the 18th birthday, Barnett and Lofaso could not reject the hypothesis that all apparent desistance was false.

Barnett, Blumstein, and Farrington (1987) then combined the approaches of Blumstein et al. (1985) and Barnett and Lofaso (1985). They analyzed conviction data collected in the CSDD, and aimed to predict the number of offenses of each person at each age and time intervals between crimes. They tested several mathematical models of criminal careers containing two key parameters: (1) p = the probability that an offender terminates his criminal career after the kth conviction; for any given offender, p is assumed to be constant for all values of k; (2) μ = the individual offending frequency per year or the annual rate at which the offender sustains convictions while free during his active career. The individual offending frequency cannot be estimated from aggregate data simply by dividing the number of convictions at each age by the number of offenders at each age, because some active offenders who have embarked on a criminal career may not be convicted at a particular age.

Barnett et al. (1987) found that models assuming that all offenders had the same p and μ did not fit the data and thus assumed that there were two categories of offenders, "frequents," and "occasionals." Each category had its own value of p and μ, which were assumed to be constant over time. They found that the model that best fitted the data had the following parameters: μ_F (conviction rate of frequents per year) = 1.14, μ_o (conviction rate of occasionals per year) = 0.41, p_F (termination probability of frequents after each conviction) = 0.10, p_o (termination probability of occasionals after each conviction) = 0.33, and α (fraction of frequents compared to occasionals) = 0.43. Thus, 43 percent of the offenders were frequents, and this group had a higher individual offending frequency and a lower probability of terminating their criminal careers after each conviction. Barnett et al. (1987) did not suggest that there were in reality only two categories of offenders, but rather that it was possible to fit the conviction data (the number of convictions of each offender at each age) using a simple model that assumed only two categories.

Barnett, Blumstein, and Farrington (1989) then carried out a test of the predictive validity of their model with the CSDD data. The model was developed on conviction data between the 10th and 25th birthdays and tested on conviction data between the 25th and 30th birthdays. The aim was to predict the number of reoffenders, the identities of reoffenders, the number of reconvictions, the age at the first reconviction, and the time intervals between reconvictions, in this follow-up period. Generally, the model performed well, but more of the frequents were reconvicted than expected, and they had more reconvictions than expected. The predictions for occasionals were excellent. For example, overall, the model predicted that 29 percent of all offenders would be reconvicted and the actual percentage was 33. For occasionals, the predicted percentage was 29 and the actual percentage was 31; for frequents, the predicted percentage was 28 but the actual percentage was 36.

It is illuminating to consider why occasionals and frequents had similar predicted reconviction probabilities. Because $p_F = 0.10$, a frequent who was convicted at age 24 had a 90 percent chance of continuing in his criminal career at age 25; because $p_o = 0.33$, an occasional who was convicted at age 24 had only a 67 percent chance of continuing in his criminal career at age 25. However, because frequents were convicted at a higher rate than the occasionals ($\mu_F = 1.14$, $\mu_o = 0.41$), a frequent who was last convicted at age 21 had only a 10 percent chance of still being an active offender at age 25, whereas an occasional who was last convicted at age 21 had almost a 30 percent chance of being an active offender at age 25. Because frequents had a higher individual offending frequency, we could be fairly certain that a long conviction-free period indicated that they had terminated their criminal careers. For occasionals, however, because they were convicted at a lower rate, a long conviction-free period did not necessarily indicate that their career had terminated. This shows how a mathematical model can help to establish convincingly when true desistance has occurred.

Using precise models as a starting point and the insights then drawn from empirical testing, Barnett et al. (1989) found that the main problem with their predictions was that a few frequents who appeared to have terminated were nevertheless unexpectedly convicted between the 25th and 30th birthdays. Therefore, they proposed that there might be some intermittency (terminating and later restarting) in criminal careers (see Piquero 2004). A few frequents seemed to cease offending at about age 19 and then were reconvicted after a period of seven to 10 years with no convictions. Barnett et al. speculated that this restarting may be connected with life changes such as losing a job or separating from a spouse, in agreement with the observed

effects of unemployment (Farrington et al. 1986) and separation (Farrington and West 1995) in the CSDD.

Challenges and Alternative Models

Michael Gottfredson and Travis Hirschi launched a series of critiques of criminal career and longitudinal research in the 1980s, including the provocatively titled paper ''The True Value of Lambda Would Appear to be Zero'' (Gottfredson and Hirschi 1986), in which they argued that individual age–crime curves were the same as the aggregate age–crime curve. Therefore, it was unnecessary to distinguish prevalence and frequency because both varied similarly with age. Blumstein, Cohen, and Farrington (1988a, 1988b) responded by arguing that the predictors and correlates of one criminal career feature (e.g., prevalence or onset) could be different from the predictors of another (e.g., frequency or desistance). They also pointed out that individual age–crime curves for frequency were very different from the aggregate age–crime curve and contended that longitudinal research was needed to test many of Gottfredson and Hirschi's key hypotheses.

Commenting on these exchanges, Greenberg (1991) proposed a more complex mathematical model that was in concordance with the Gottfredson–Hirschi theory. Instead of assuming two categories of offenders, he proposed that there was a continuous (gamma) distribution of λ over offenders and showed that his model could fit the frequency distribution of number of offenses over persons in four studies. Barnett et al. (1992) pointed out that Greenberg's model only fitted the frequency distribution of offenses over persons, whereas Barnett et al. (1989) predicted several criminal career features including the fraction who recidivated, the identity of those who recidivated, the total number of reconvictions, the age of the first reconviction, and the length of interconviction intervals. They also argued that Greenberg's model allowed no termination of offending, whereas Barnett et al. (1987, 1989) concluded that crime decreased with age because of termination (since the interconviction intervals did not increase with age). Greenberg (1992) replied that the fit of his models was not affected by assuming that λ was constant or that λ decreased with age.

Another latent propensity model was proposed by Rowe, Osgood, and Nicewander (1990), who set out to demonstrate that their latent trait approach could provide a parsimonious model that unified knowledge about criminal careers. Starting with the Gottfredson and Hirschi-friendly assumption that an underlying trait of criminal propensity or self-control is sufficient to explain criminal offending, they characterized the criminal career

paradigm as being based on just four parameters (participation, frequency, duration, and specialization) and that it assumed that the population is necessarily divided into two distinct categories of offenders and nonoffenders. Their latent trait approach proposed that criminal propensity, a trait which is not directly observable, and is assumed to be distributed lognormally over the entire population, was sufficient to explain both participation in and frequency of offending.

Rowe et al. (1990) assumed that individuals in the population occupied a relatively stable position on the propensity scale. The normally distributed propensity Θ was transformed into an individual offending rate according to the equation $\lambda = \exp[a^*(\Theta - b)]$; the parameters a and b were chosen by maximum likelihood fitting to known, whole population, aggregate offending rate data (the proportion of the population with 0, 1, 2, … 9+ offenses per year). Like Greenberg (1991), they showed that their model fitted the frequency distribution of offenses over persons in four studies. Rowe et al. went on to fit their model to a variety of subpopulations including gender and race categories and offense types and claimed plausible fits in all cases. They concluded that the ability of the single latent trait model to fit this variety of data showed that the criminal career models that distinguished participation and frequency were unnecessary and implausible. Although still proponents of the latent trait approach to explaining offending, Osgood and Rowe (1994) aimed to " … build bridges between theoretical criminology, the study of criminal careers, and policy-relevant research" (*p.* 517). The main thrust of their paper described a variety of linear models relating various types of explanatory variables to observed outcomes.

The models developed in the above papers assume a distribution of λ from which estimates of participation and aggregate frequency can be derived. With the addition of the linear modeling, these elements of criminal careers might be derivable from measurable individual factors. The parameters of the models are empirically determined to fit the data but they are not related directly to real-world factors which might be amenable to policy interventions. These models provide no insights into many other criminal career features, such as onset, duration, recidivism, residual career duration, desistance, and, most importantly, the relationship of age with any of these.

Nagin and Land (1993) then proposed a more general model that does not make specific predictions that are as extensive as others described here. Like Greenberg (1991) and Rowe et al. (1990), they assumed that λ varies continuously over the population and that it was unnecessary to distinguish

offenders from nonoffenders or study onset and desistance. Departing from the models above, however, they made no assumptions about the form of the λ distribution and assumed that λ was a function of age, gender, time-stable characteristics of the individual and persistent unobserved heterogeneity. They fitted conviction data from the CSDD and also proposed that there were three categories of offenders: high-rate chronics, low-rate chronics, and adolescent-limited offenders. Nagin, Farrington, and Moffitt (1995) then analyzed the individual characteristics, behaviors, and social circumstances of these groups in the CSDD between ages 10 and 32.

A key assumption of Nagin and Land's group-based trajectory model is that it does not take as given that the aggregate age–crime curve applies equally to all persons, and most studies using their method have identified important individual variability in longitudinal offending patterns (Jennings and Reingle 2012; Piquero 2008). At the same time, while researchers have used this technique to identify categories of individuals with different numbers of offenses at each age, they have not investigated other features of criminal careers. It is also difficult to assign individuals to trajectories unambiguously, and it is sometimes unclear whether trajectories differ in degree or in kind. The value of trajectory modeling is quite controversial, as shown by the exchange between Skardhamar (2010) and Brame, Paternoster, and Piquero (2012). Skardhamar argued that, even when no groups exist in reality and the data are truly continuous, trajectory modeling will yield several categories of offenders. In response, Brame et al. did not disagree with this argument but contended that trajectory modeling was useful for description and visualization of criminal career data and for testing theories.

While the trajectory modeling can provide some useful insights, there is a more fundamental objection to its use as a method of criminal career modeling. The trajectories of the groups identified are invariably modeled using quadratic or cubic equations in age. While these models often provide a good fit to the data, the models themselves are atheoretical (Sampson and Laub 2005). The parameters of the models have little meaning in the real world and may not be closely correlated with attributes of offenders, and the modeling itself provides little direct insight into the causal processes involved in the progression of criminal careers. Yet, trajectory modeling does permit researchers to investigate how risk and protective factors are related to different trajectory groups (Nagin and Tremblay 2001, 2005). While disagreements about the utility of trajectory models have been a prominent part of the discussion of criminal careers in recent years, some researchers have reinvigorated aspects of earlier criminal career models.

Recent Research by MacLeod et al.

Surprisingly, since the 1980s, there have been very few attempts to develop and test predictive criminal career models of the type developed by Blumstein and his colleagues. An exception is the book, *Explaining Criminal Careers*, by MacLeod et al. (2012) that attempted to resuscitate mathematical models of criminal careers. This book went beyond previous research in a number of ways. First, different categories of offenders were discovered using graphical methods. Second, the analyses were based on conviction records of very large U.K. representative samples, including comparisons across birth cohorts and time periods. Third, the individual offending frequency and the probability of recidivism were decoupled, so that it was possible to have a category of offenders with a low frequency and a high probability of reoffending. Fourth, models were proposed to explain the onset process and the complete age–crime curve, not just persistence and desistance of offending.

MacLeod et al. first tried to predict the growth in recidivism probabilities after each successive offense. They found that the data could be fitted very well by assuming that there were two categories of offenders with constant but different probabilities of recidivating after each conviction. For the 1953 national U.K. birth cohort followed up to 1999, these were .84 and .35, respectively, for males and .81 and .19, respectively, for females; 27 percent of males and 9 percent of females were in the high-probability category. They then tried to predict the distribution of times between successive offenses. Again, the data could be fitted very well by assuming that there were two categories of offenders, in this case with constant but different frequencies of offending. For the 1953 national U.K. birth cohort, these were .85 and .21 convictions per year, respectively, for males, and .97 and .23 convictions per year, respectively, for females; 56 percent of males and 54 percent of females were in the high-frequency category.

MacLeod et al. then fitted the frequency distribution of the number of convictions of each offender by assuming that there were three categories of offenders: high frequency–high probability, low frequency–low probability, and low frequency–high probability. In contrast to the trajectory method, this assumes the existence of specific categories a priori, makes explicit predictions about them, and then tests the assumptions against data. The proportions in the three categories were 19 percent, 73 percent, and 8 percent, respectively, for males and 7 percent, 91 percent, and 2 percent, respectively, for females. Interestingly, Piquero, Sullivan, and Farrington

(2010) had previously distinguished short-term high-frequency offenders from long-term low-frequency offenders in the CSDD.

MacLeod et al. then showed that it was possible to fit the age–crime curve (the aggregate number of convictions at each age). Their previous assumptions fitted the down-slope of the curve after the peak at age 17 very well, but it was necessary to add new assumptions to predict the up-slope before the peak. They fitted the age at first conviction assuming a proportional hazard survival process modified by an ogive function modeling the increasing probability of being convicted after an offense from age 10 to age 17. This added two parameters, one for the slope of this increase (.54) and one for the age when the probability was .5 (14.7). With some further mathematics, MacLeod et al. derived predictions for ages at first conviction and ages at all convictions, making accurate predictions (within 1 percent) of the number of offenders reconvicted within 15 months, for a completely new sample of adult offenders convicted in 2004. The models predicted that, of 4,833 convicted offenders, 2,708 would be reconvicted; the actual number was 2,681.

Returning to the theme of this article, MacLeod et al. showed that several features of criminal careers can be quantitatively predicted on the basis of very simple assumptions. The fit of models to the data was very good; for example, MacLeod et al. (2012:36) reported a correlation of $r = .9999$ between the model predictions and the distribution of times between reconvictions. Some assumptions remain controversial, such as the argument that changes with age in the aggregate crime rate are generally caused by compositional changes (e.g., in the prevalence of offenders) not by changes with age in the individual offending frequency.

Conclusions

The empirical study of criminal careers has been informed greatly by the development of various criminal career models. The ideal is to propose simple models that logically lead to a large number of testable predictions of criminal career results. We are not convinced that the more complex models that have been proposed since the 1980s (e.g., assuming continuous distributions of λ) are superior to the simple models in the range of results that they can predict. We believe in the principle of Occam's razor: no more assumptions should be made than are necessary.

A key issue is: In predicting criminal careers, is it more useful to focus on features such as prevalence, frequency, onset, termination, and duration, or is it more useful to assume that all these features (and others such as

specialization and escalation) are generated by an underlying continuous dimension of criminal propensity, so that it is unnecessary even to distinguish between offenders and nonoffenders? We believe that there is more value, in explaining and predicting offending, to focus on criminal career features. We also believe that this is more useful for criminal justice policies (DeLisi and Piquero 2011; Piquero 2011).

The time is ripe to build on simple models of the age–crime curve to explain a wider range of criminological findings. Parameters can be added to specify quantitatively how particular risk and protective factors influence the onset, frequency, persistence, termination, escalation, and deescalation of criminal careers at different ages. Recall that Blumstein et al. (1985) showed how a risk factor score predicted the recidivism processes of persisters and desisters. The models could be expanded to focus on convictions, arrests, or self-reports; different types of crimes, specialization, and versatility; gender and racial/ethnic differences; geographical differences between U.S. states and between countries; different time periods and birth cohorts; and so on. The criminal career models of the 1980s hold the promise of converting criminology into a more predictive and accurate science. Developing and testing different mathematical models of criminal careers can yield a better understanding of criminal behavior over the life course and, in turn, help to develop more effective prevention and intervention strategies.

Declaration of Conflicting Interests

Funding

References

Atkinson, Richard C., Gordon H. Bower, and Edward J. Crothers. 1965. *An Introduction to Mathematical Learning Theory*. London, UK: Wiley.

Barnett, Arnold, Alfred Blumstein, Jacqueline Cohen, and David P. Farrington. 1992. "Not All Criminal Career Models Are Equally Valid." *Criminology* 30: 133-40.

Barnett, Arnold, Alfred Blumstein, and David P. Farrington. 1987. "Probabilistic Models of Youthful Criminal Careers." *Criminology* 25:83-107.

Barnett, Arnold, Alfred Blumstein, and David P. Farrington. 1989. "A Prospective Test of a Criminal Career Model." *Criminology* 27:373-88.

Barnett, Arnold and Anthony J. Lofaso. 1985. "Selective Incapacitation and the Philadelphia Cohort Data." *Journal of Quantitative Criminology* 1:3-36.

Blumstein, Alfred and Jacqueline Cohen. 1979. "Estimation of Individual Crime Rates from Arrest Records." *Journal of Criminal Law and Criminology* 70: 561-85.

Blumstein, Alfred, Jacqueline Cohen, and David P. Farrington. 1988a. "Criminal Career Research: Its Value for Criminology." *Criminology* 26:1-35.

Blumstein, Alfred, Jacqueline Cohen, and David P. Farrington. 1988b. "Longitudinal and Criminal Career Research: Further Clarifications." *Criminology* 26:57-74.

Blumstein, Alfred, Jacqueline Cohen, and Daniel S. Nagin, eds. 1978. *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*. Washington, DC: National Academy Press.

Blumstein, Alfred, Jacqueline Cohen, Jeffrey A. Roth, and Christy A. Visher, eds 1986. *Criminal Careers and "Career Criminals."* Vol. 2. Washington, DC: National Academy Press.

Blumstein, Alfred, David P. Farrington, and Soumyo Moitra. 1985. "Delinquency Careers: Innocents, Desisters, and Persisters." Pp. 187-222 in *Crime and Justice*, Vol. 6, edited by Michael Tonry and Norval Morris. Chicago: University of Chicago Press.

Blumstein, Alfred and Soumyo Moitra. 1980. "The Identification of "Career Criminals" from "Chronic Offenders" in a Cohort." *Law and Policy Quarterly* 2: 321-34.

Brame, Robert, Raymond Paternoster, and Alex R. Piquero. 2012. "Thoughts on the Analysis of Group-based Developmental Trajectories in Criminology." *Justice Quarterly* 29:469-90.

Carr-Hill, G. A. and Roy Carr-Hill. 1972. "Reconviction as a Process." *British Journal of Criminology* 12:35-43.

DeLisi, Matt and Alex R. Piquero. 2011. "New Frontiers in Criminal Careers Research, 2000-2011: A State-of-the-art Review." *Journal of Criminal Justice* 39:289-301.

Farrington, David P., Jeremy W. Coid, Louise Harnett, Darrick Jolliffe, Nadine Soteriou, Richard Turner, and Donald J. West. 2006. *Criminal Careers up to Age 50 and Life Success up to Age 48: New Findings from the Cambridge Study in Delinquent Development*. Research Study No. 299. London, UK: Home Office.

Farrington, David P., Bernard Gallagher, Lynda Morley, Raymond J. St Ledger, and Donald J. West. 1986. "Unemployment, School Leaving and Crime." *British Journal of Criminology* 26:335-56.

Farrington, David P., Alex R. Piquero, and Wesley G. Jennings. 2013. *Offending from Childhood to Late Middle Age: Recent Results from the Cambridge Study in Delinquent Development*. New York: Springer.

Farrington, David P., Maria M. Ttofi, Rebecca V. Crago, and Jeremy W. Coid. 2014. ''Prevalence, Frequency, Onset, Desistance and Criminal Career Duration in Self-reports Compared with Official Records.'' *Criminal Behaviour and Mental Health* 24:241-53.

Farrington, David P. and Donald J. West. 1993. ''Criminal, Penal and Life Histories of Chronic Offenders: Risk and Protective Factors and Early Identification.'' *Criminal Behaviour and Mental Health* 3:492-523.

Farrington, David P. and Donald J. West. 1995. ''Effects of Marriage, Separation and Children on Offending by Adult Males.'' Pp. 249-81 in *Current Perspectives on Aging and the Life Cycle, Vol. 4: Delinquency and Disrepute in the Life Course*, edited by John Hagan. Greenwich, CT: JAI Press.

Gottfredson, Michael R. and Travis Hirschi. 1986. ''The True Value of Lambda Would Appear to be Zero: An Essay on Career Criminals, Criminal Careers, Selective Incapacitation, Cohort Studies, and Related Topics.'' *Criminology* 24:213-34.

Greenberg, David F. 1979. *Mathematical Criminology*. New Brunswick, NJ: Rutgers University Press.

Greenberg, David F. 1991. ''Modeling Criminal Careers.'' *Criminology* 29:17-46.

Greenberg, David F. 1992. ''Comparing Criminal Career Models.'' *Criminology* 30: 141-47.

Greenwood, Peter W. and Allan Abrahamse. 1982. *Selective Incapacitation*. Santa Monica, CA: Rand.

Jennings, Wesley G. and Jennifer M. Reingle. 2012. ''On the Number and Shape of Developmental/Life-course Trajectories: A State-of-the-art Review.'' *Journal of Criminal Justice* 40:472-89.

Laming, Donald. 1973. *Mathematical Psychology*. London, UK: Academic Press.

MacLeod, John F., Peter G. Grove, and David P. Farrington. 2012. *Explaining Criminal Careers: Implications for Justice Policy*. Oxford, UK: Oxford University Press.

Moffitt, Terrie E. 1993. ''Adolescence-limited and Life-course-persistent Antisocial Behavior: A Developmental Taxonomy.'' *Psychological Review* 100: 674-701.

Moffitt, Terrie E. 2006. ''Life-course-persistent versus Adolescence-limited Antisocial Behavior.'' Pp. 570-98 in *Developmental Psychopathology, Vol. 3: Risk, Disorder, and Adaptation*, edited by Dante Cicchetti and Donald J. Cohen. New York: Wiley.

Nagin, Daniel S. 2005. *Group-based Modeling of Development*. Cambridge, MA: Harvard University Press.

Nagin, Daniel S., David P. Farrington, and Terrie E. Moffitt. 1995. ''Life-course Trajectories of Different Types of Offenders.'' *Criminology* 33:111-39.

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.

Nagin, Daniel S. and Richard E. Tremblay. 2001. "Parental and Early Childhood Predictors of Persistent Physical Aggression in Boys from Kindergarten to High School." *Archives of General Psychiatry* 58:389-94.

Nagin, Daniel S. and Richard E. Tremblay. 2005. "Developmental Trajectory Groups: Fact or Fiction?" *Criminology* 43:873-904.

Osgood, D. Wayne and David C. Rowe. 1994. "Bridging Criminal Careers, Theory, and Policy through Latent Variable Models of Individual Offending." *Criminology* 32:517-54.

Piquero, Alex R. 2004. "Somewhere Between Persistence and Desistance: The Intermittency of Criminal Careers." Pp. 102-28 in *After Crime and Punishment: Pathways to Offender Reintegration*, edited by Shadd Maruna and Russ Immarigeon. London, UK: Routledge.

Piquero, Alex R. 2008. "Taking Stock of Developmental Trajectories of Criminal Activity over the Life Course." Pp. 23-78 in *The Long View of Crime: A Synthesis of Longitudinal Research*, edited by Akiva Liberman. New York: Springer.

Piquero, Alex R. 2011. "Invited Address: James Joyce, Alice in Wonderland, the Rolling Stones, and Criminal Careers." *Journal of Youth and Adolescence* 40: 761-75.

Piquero, Alex R., David P. Farrington, and Alfred Blumstein. 2003. "The Criminal Career Paradigm." Pp. 359-506 in *Crime and Justice*, Vol. 30, edited by Michael Tonry. Chicago: University of Chicago Press.

Piquero, Alex R., David P. Farrington, and Alfred Blumstein. 2007. *Key Issues in Criminal Career Research: New Analyses of the Cambridge Study in Delinquent Development*. Cambridge, UK: Cambridge University Press.

Piquero, Alex R., Christopher J. Sullivan, and David P. Farrington. 2010. "Assessing Differences between Short-term, High-rate Offenders and Long-term, Low-rate Offenders." *Criminal Justice and Behavior* 37:1309-29.

Piquero, Alex R. and David Weisburd, eds. 2010. *Handbook of Quantitative Criminology*. New York: Springer.

Rowe, David C., D. Wayne Osgood, and W. Alan Nicewander. 1990. "A Latent Trait Approach to Unifying Criminal Careers." *Criminology* 28:237-70.

Sampson, Robert J. and John H. Laub. 2005. "Seductions of Method: Rejoinder to Nagin and Tremblay's "Developmental Trajectory Groups: Fact or Fiction?"" *Criminology* 43:905-13.

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

Skardhamar, Torbjorn. 2010. "Distinguishing Facts and Artefacts in Group-based Modeling." *Criminology* 48:295-320.

Visher, Christy A. 1986. "The Rand Inmate Survey: A Reanalysis." Pp. 161-211 in *Criminal Careers and "Career Criminals*," vol. 2, edited by Alfred Blumstein, Jacqueline Cohen, Jeffrey A. Roth, and Christy A. Visher. Washington, DC: National Academy Press.

Wolfgang, Marvin E., Robert M. Figlio, and Thorsten Sellin. 1972. *Delinquency in a Birth Cohort*. Chicago: University of Chicago Press.

Author Biographies

David P. Farrington is Emeritus Professor of Psychological Criminology in the Institute of Criminology, Cambridge University. He received the Stockholm Prize in Criminology in 2013, and he is Chair of the American Society of Criminology Division of Developmental and Life-Course Criminology. In addition to over 650 published journal articles and book chapters on criminological and psychological topics, he has published nearly 100 books, monographs and government reports.

John F. MacLeod is a Visiting Scholar at the Institute of Criminology, Cambridge University, where he continues his research interests in criminal careers, the causes and correlates of crime, and the use of mathematical modeling. He retired in 2007 from the UK Home Office, where for over 20 years he undertook research, analysis and modeling in support of criminal justice policy. He is a Fellow of the Institute of Mathematics and its Applications and a Chartered Mathematician.

Alex R. Piquero is Ashbel Smith Professor of Criminology and Associate Dean of Graduate Programs in the School of Economic, Political and Policy Sciences at the University of Texas at Dallas. His research interests include criminal careers, criminological theory, and quantitative research methods. He has received several research, teaching, and service awards and is a Fellow of both the American Society of Criminology and the Academy of Criminal Justice Sciences. In 2014, he received the University of Texas Regents' Outstanding Teaching Award.