1 Looking Forward

Criminologists generally agree that age is one of the strongest predictors of criminal activity and that a disproportionately large share of offending is committed by those who are in the age range between mid-adolescence and young adulthood (Farrington 1986). In fact, the relationship between age and crime is one of the most robust empirical findings in criminology, or as Hirschi and Gottfredson (1983: 552) state, ‘this distribution thus represents one of the brute facts of criminology.’

When criminologists speak of the relationship between age and crime, they usually are referring to the ‘age–crime curve’. The aggregate age–crime curve (which is computed by dividing the total number of arrests of individuals of a given age by the total population size of the specific age) indicates: (1) a sharp increase in the arrest rate in the early teen years; (2) a peak age of arrest in the late teen or early young adult years (depending on the crime type); and (3) a decrease in the rate of arrest over the remaining age distribution. Graphically, the distribution of arrests over the age range resembles the lognormal or gamma probability density functions, distributions characterized by both a peak and a long right tail (see Britt 1992).

Consider, for example, the two panels in Figure 1.1 that contain the aggregate age–crime curves for both violent and property FBI Index crimes in 1980, 1994, and 2000.1 Although there is some parametric invariance in these age–crime curves (i.e. the mean, mode, skew, and kurtosis are not identical in each year), the general age–crime relationship described above can be readily seen. In each of the panels and for each year, crime rates dramatically increase in

---

1 Violent Index crimes include murder and non-negligent manslaughter, forcible rape, aggravated assault, and robbery. Property Index crimes include burglary, larceny theft, motor vehicle theft, and arson.
Figure 1.1. Violent and property aggregate age–crime curves in the United States in 1980, 1994, and 2000.

early adolescence, peak in late adolescence or early adulthood, and then continually decrease over the remaining age distribution. Further evidence of the relationship between age and crime can be found in studies that analyze data relating crime rates to aggregates of various sizes. These studies consistently report that, overall, the age distribution of any population is inversely related to its crime rate (Cohen and Land 1987; Hirschi and Gottfredson 1983; Steffensmeier and Harer 1987; Steffensmeier et al. 1989).

Beyond that basic description, however, the relationship between age and crime is the fundamental source of many controversies in criminology, controversies that have sometimes led to rather rancorous debates between researchers. According to Lauritsen (1998: 127), ‘Few substantive issues in criminology have been more contentious than those raised by the study of age and crime. While most social scientists agree that the aggregate age–crime curve reaches a peak during late adolescence and declines rapidly thereafter, there are ongoing debates about the theoretical meaning of this “brute fact”.’

This study, which employs the use of three samples of serious chronic youthful offenders, examines three key questions regarding the relationship between age and crime: (1) are there two (or more) discrete groups of offenders with distinct age–crime curves concealed within the aggregate age–crime curve; (2) how stable are individual differences in the propensity to commit criminal acts across the life course; and (3) does the relationship between offending at one time period and a subsequent period endure when time-stable individual differences in criminal propensity are controlled?

The ‘Great Debate’ Concerning the Age–Crime Curve

Beginning in the mid-1980s, the discipline of criminology witnessed what Vold, Bernard, and Snipes (1998: 285) have called the ‘Great Debate’ concerning the relationship between age and crime. This debate involved a dispute over whether one finds the same relationship between age and crime with individual-level data as that which is observed when analyzing aggregate data. Two main factions formed within this debate—one represented by Hirschi and Gottfredson (1983; Gottfredson and Hirschi 1986, 1988, 1990) and the other by Blumstein and his colleagues (Blumstein and

Hirschi and Gottfredson (1983) contend crime is everywhere inversely related to age at both the individual and aggregate levels of analysis. Thus, the relationship between age and crime is deemed to be invariant; all people, everywhere, within any historical period (and with both official statistics and self-report data), tend to commit less crime as they age regardless of offense type. Hirschi and Gottfredson argue that age-specific offense rates increase dramatically from age 10 until age 17, and then continually decrease thereafter. In addition, Hirschi and Gottfredson (1983: 565, 578–9) emphasize that the decrease in offending with age occurs regardless of the offender’s criminal propensity (i.e. no matter whether the individual’s criminality is high or low). Thus, they expect that (after the peak years) the rate of offending will decrease with age, even among those serious and/or chronic offenders who are still criminally active.

Blumstein and his colleagues, on the other hand, argue that age is not inversely related to criminal offending at the individual level of analysis among active offenders. Blumstein and his colleagues concede that both participation in criminal activity and the incidence rates of offending vary inversely with age at the population level. However, they contend that Gottfredson and Hirschi confuse changes in participation and incidence rates with changes in the frequency of individual offending among active offenders (often referred to as lambda). While Gottfredson and Hirschi argue that incidence rates decline because there is a decrease in frequency of offending by active offenders, Blumstein et al. argue the incidence rate declines because there are fewer active offenders as age increases. Thus, it is the effect of offenders beginning (onset) and terminating (desisting) their criminal careers that is largely driving the shape of the age–crime curve. In short, Blumstein and his colleagues argue that as long as offenders are active, they will continue to commit crimes at a relatively constant rate independent of their age. If this is true, it has profound implications for crime control policies as the incapacitation of active offenders would significantly reduce the crime rate.

2 Piquero, Farrington, and Blumstein (2003) present an excellent and detailed history of the criminal careers approach to explaining crime patterns (and research based on this approach), as well as detailing the debate between Blumstein and colleagues and Gottfredson and Hirschi.
Certainly, one of the major points that Blumstein and his colleagues are trying to convey is that the shape of the age–crime curve could be the result of a process other than offenders simply committing fewer crimes as they age, and thus, that caution is imperative when offering explanations for the empirical shape of the age–crime curve. More specifically, they indicate that the age–crime curve is driven by two processes: participation rates and incidence rates. A change in either one of these rates affects the empirical shape of the age–crime curve. Still, their argument, which has been made repeatedly, is that as long as offenders are active, they will continue to commit crimes at a relatively constant rate independent of their age. As Farrington (1986: 218) notes, ‘they [Blumstein and his colleagues] have consistently argued that the individual crime rate or incidence of offending [lambda] is constant during a criminal career and that changes in aggregate crime rates reflect changes in prevalence.’ In addition, Farrington (1986: 189) himself argues, ‘age–crime curves for individuals do not resemble the aggregate curve since incidence [lambda] does not change consistently between the onset and termination of criminal careers.’ Indeed, recently Farrington (1997: 365) argued that a ‘30 year old offender commits offenses at roughly the same rate as an 18 year old offender, although offenders are more prevalent in the population of 18 year olds than in the population of 30 year olds.’

Brame and Piquero (2003: 108) recently noted, ‘a key issue is whether individual age–crime curves vary primarily in terms of the overall volume of offending or whether the aggregate age–crime curve is a mixture of micro-level curves that vary widely in shape,’ while Farrington (2003: 226) noted in his discussion of contentious issues facing developmental and life-course criminology that information is lacking with respect to ‘how much there are distinctly different behavioral trajectories.’ In other words, the primary source of contention between various researchers still concerns the causes of the inverted-J shape of the age–crime curve, but the specific disagreement has shifted somewhat from purely focusing on whether the relationship between age and crime is constant among ‘active offenders.’

Recent issues related to the relationship between age and crime focus on three key questions: (1) how stable or unstable are individual differences in criminal behaviour across the

---

3 Chapter 3 reviews the evidence from longitudinal studies of criminal offending patterns over time.
life course; (2) are there two discrete groups of offenders in the offender population, each with their own age–crime curves that differ from the overall aggregate curve, but which when aggregated together produce the observed overall curve; and (3) is there a significant relationship between criminal activity at adjacent ages (or periods) after controlling for persistent differences in the propensity to offend? Although these issues are sometimes treated as mutually exclusive, they are actually highly interconnected and can all be viewed within the bounds of the long-standing ‘paradox of persistence’ phenomenon in criminology (Cohen and Vila 1996).

The Paradox of Persistence

The ‘paradox of persistence’ refers to the consistent finding that when looking in reverse (or retrospectively) researchers find that most adult criminal offenders were juvenile delinquents, yet when looking forward (or prospectively) in the lives of juvenile delinquents researchers find that most delinquents do not go on to become adult criminal offenders (see e.g. Blumstein et al. 1986; Caspi and Moffitt 1992; Cernkovich and Giordano 2001; Cline 1980; Gove 1985; Loeber and Le Blanc 1990; McCord 1980; Robins 1978; Sampson 2000; Sampson and Laub 1993, 1997; Tracy and Kempf-Leonard 1996). An oft-cited quotation from Robins (1978: 611) perhaps best summarizes the issue: ‘adult antisocial behavior virtually requires childhood antisocial behavior, yet most antisocial children do not become antisocial adults.’ Thus, on the one hand, there is a considerable amount of continuity in behavior over time, evidenced by the fact that few, if any, criminologists would argue with the statement that the presence (and/or frequency) of delinquent criminal activity during childhood and adolescence is one of the best, if not the best, predictor of adult criminality. Yet, at the same time, the relationship between juvenile and adult criminal activity is not a deterministic one, and a number of juvenile offenders are able to escape the criminal lifestyle and do not become ‘career criminals.’ In other words, there is both continuity and change (discontinuity) over time, and while this is a relatively straightforward (and some might argue simple) statement, the etiological explanation of this relationship is actually very complex. This relationship, in fact, is the source of a key theoretical controversy in the discipline of criminology (Cohen and Vila 1996): how can
one simultaneously explain the sources of both continuity and change in criminal behavior over time? Any valid explanation of the sources of continuity and change (discontinuity) in criminal offending patterns must consequentially produce an explanation that is consistent with the observed shape of the age–crime curve and account for whether offending at adjacent ages is causally related after accounting for individual differences in the propensity to offend.

**Explaining the Paradox of Persistence**

Three broad theoretical explanations have been proffered to explain both the paradox of persistence and the shape of the age–crime curve: (1) population heterogeneity, (2) state dependence, and (3) the dual taxonomy approaches. According to the population heterogeneity explanation (in its purest form), continuity and change in criminal offending patterns over time are explained entirely by *time-invariant differences* in a latent proneness to engage in criminal activity. Population heterogeneity theories are sometimes called *latent trait* theories because they posit that there is variation across the population on a persistent, underlying or *latent* variable that explains crime (Nagin and Farrington 1992a, 1992b; Nagin and Paternoster 1991). This latent variable is argued to be either unmeasured or poorly measured (Cohen and Vila 1996). According to the population heterogeneity position, all offenders follow the same age–crime curve and all offenders are argued to decrease their offending over time. The individuals who have higher levels of the latent variable, however, will engage in criminal activity earlier in life, persist in committing criminal acts further into adulthood, and commit criminal offenses at a higher rate at all points in time. Individuals lower on the distribution of latent propensity, on the other hand, are argued to begin offending later in life, end their offending earlier in the life course, and commit offenses at a lower rate at all points in time. In other words, continuity and change is explained entirely by between-individual differences in latent criminal propensity. The association between criminal activity at any two points in time (e.g. two adjacent ages) is argued to be caused by the latent propensity, and controlling for the underlying differences in propensity eliminates any relationship between crime at any two points in time. Gottfredson and Hirschi’s (1990)
self-control theory is arguably the leading population heterogeneity theory.

A pure state dependence explanation, on the other hand, argues that the *propensity to engage in crime is malleable* over the age distribution. Continuity in criminal activity arises as a result of the negative cumulative consequences of earlier criminal activity and/or continued engulfment in the ‘criminal lifestyle’, whereas change results from experiencing positive events that can potentially mitigate one’s criminal propensity (e.g. getting married or obtaining a good job). The shape of the age–crime curve is derived from the fact that criminal propensity is variable over the life course and is at its highest levels in the mid- to late adolescent years. An important proposition of state dependence arguments is that criminal activity at one point in time is *causally* related to criminal activity at a later point in time. The age-graded theory of informal social control proposed by Sampson and Laub (1993) is often considered as an example of a state dependence theory.4

A third primary theoretical perspective, the dual taxonomy explanation, argues that the criminal offender population is comprised of two empirically distinct offender categories, each with its own etiological explanation. In dual taxonomy approaches, there is a larger group of offenders who only engage in criminal activity during adolescence, while the second smaller group is ‘life-course persistent’ in their criminal activity. Thus, since the aggregate age–crime curve mixes these two groups together, it takes on its observed shape. The change is the result of desistence by the ‘adolescent-limited’ offenders, whereas continuity results from the continuing criminal activity of the ‘life-course-persistent’ group (Moffitt 1993, 1997). According to this stream of theoretical thought, the association between criminal activity at two points in time is argued to be spurious for the life-course-persistent group, and causally related for the adolescent-limited group. The dual taxonomy theory of Moffitt (1993, 1997) is an example of this type of theoretical approach.

4 As described in the next chapter, the age-graded theory of Sampson and Laub (1993) is better described as a ‘mixed’ theory that allows for both state dependence and enduring individual differences (Nagin and Paternoster 2000). The primary assumption of this theory, however, is that criminal propensity can be changed as a result of changing levels of social control over the life course, and thus it is generally referred to as a theory of state dependence.
Overview of the Study

As noted above, this study examines three key, often controversial questions. These questions will be examined using arrest data from the serious youthful offender population, a population of offenders that is rarely included in examinations of the issues addressed in this study. To date, research has largely ignored the empirical question of whether or not serious youthful offenders are a homogeneous group or a heterogeneous bunch of groups with differential rates and trajectories of criminal activity across the age distribution.\(^5\)

The remainder of this study will proceed in the following manner. In Chapter 2, we present a more comprehensive description of the population heterogeneity, state dependence, and dual taxonomy perspectives, with special emphasis on the self-control theory of Gottfredson and Hirschi (1990), the age-graded theory of informal social control proposed by Sampson and Laub (1993), and the dual taxonomy theory of Moffitt (1993).

In Chapter 3, we review the extant empirical literature on the topics of concern in this study. The specific hypotheses examined in this study are also presented in Chapter 3. In Chapter 4, we describe the California Youth Authority (the source of the samples used herein), the state agency responsible for housing the most seriously delinquent/criminal youthful offenders in that state.

The data and statistical methods employed in this study are described in Chapter 5. More specifically, the data utilized herein are three samples of individuals released on parole from the California Youth Authority in fiscal years 1981–2, 1986–7, and 1991–2. The sources of data and variables used in the analyses are described in this chapter. Chapter 5 concludes with a brief

---

\(^5\) Previous research too often treats this group of serious offenders as a homogeneous group (especially in terms of the continuity of their behavioral patterns) because they only comprise a very small fraction (<6%) of general population samples (e.g. birth cohort samples, school samples). They are usually treated as a homogeneous group via a binary variable (usually labelled ‘chronic’, ‘serious’, and/or ‘persistent’ offenders on the basis of making some often ex post facto cut-off point) and then are compared against the low-rate offenders and non-offenders. Sampson (2000) criticizes this static, between-individual research (which produces conclusions indicating high-rate offenders have elevated risks of continuity in offending relative to non-chronic offenders) because it tends to reify the fixed categories and neglects the nature and extent of within-individual change.
Looking Forward

description of the statistical methods employed in this study, namely the finite mixture or semiparametric random effects models of Nagin and Land (1993; Land and Nagin 1996; Land, McCall, and Nagin 1996; Nagin 1999), as well as parametric random effects panel methods.

We present a descriptive summary of the three data sets in Chapter 6. This chapter includes a description of the characteristics of the cases (e.g. ethnicity, gang membership, and drug abuse), the age at first criminal arrest, the types of offenses the individuals in the samples were arrested for perpetrating, their adult incarceration experiences, and the mortality patterns of these individuals.

The results of the substantive analyses are presented in Chapters 7 and 8. Chapter 7 presents the results from applications of the Nagin and Land (1993) finite mixture model to each of the three samples. The results in Chapter 7 are, in turn, used to investigate the age–crime curves for distinct ‘latent classes’ of offenders who share a similar age-specific offending trajectory, with an emphasis on examining whether the relationship between age and crime is invariant across the latent classes. Also, the latent classes derived from the application of the finite mixture models in this chapter will form the basis for subsequent analyses in Chapter 8 that test whether there is a relationship of past to subsequent criminal activity after controlling for unobserved or ‘hidden heterogeneity’ (Land and Nagin 1996).

The methodological approach used in Chapter 8 is the multi-method approach described by Bushway, Brame, and Paternoster (1999), who recommend using several different statistical models (each with different assumptions) to estimate the relationship between past and subsequent criminal behaviour. This is critical because recent research suggests that conclusions arrived at in some previous empirical investigations of this issue could possibly be method-specific, thereby bringing into question the robustness and reliability of the prior findings. To the degree that one can robustly replicate the findings using different methods that have different assumptions, one can be more assured of the existence of the estimated effect (Bushway, Brame, and Paternoster 1999). Similarly, replicating findings across multiple samples would lend further support to the robustness of any observed effect.
The final chapter, Chapter 9, discusses the main findings of this study as evidence in support of or contrary to the hypotheses described in Chapter 3. Chapter 9 then concludes with a discussion of the limitations of this study, policy implications, and directions for future research concerning the topic of continuity and discontinuity in criminal offending patterns.