The author(s) shown below used Federal funds provided by the U.S. Department of Justice and prepared the following final report:

Document Title:	**Can a Leopard Change its Spots? Continuity and Change in Criminal Offending Patterns among Three Samples of Serious Chronic Offenders**
Author(s):	**Michael E. Ezell ; Lawrence E. Cohen**
Document No.:	**213005**
Date Received:	**January 2006**
Award Number:	**98-CE-VX-0026**

This report has not been published by the U.S. Department of Justice. To provide better customer service, NCJRS has made this Federally-funded grant final report available electronically in addition to traditional paper copies.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Can a Leopard Change its Spots?

Continuity and Change in Criminal Offending

Patterns among Three Samples of Serious Chronic Offenders

Do not cite or quote without the permission of the authors

Michael E. Ezell
Department of Sociology, Vanderbilt University

Lawrence E. Cohen
Department of Sociology, University of California at Davis

October 25, 2002

Final Report
National Institute of Justice
Grant #98-CE-VX-0026

ABSTRACT

One of the foremost controversies in contemporary theoretical criminology concerns how to simultaneously explain the existence of both continuity and discontinuity in patterns of criminal behavior over time. Three broad theoretical paradigms have been used to explain both continuity (stability) and discontinuity (change): population heterogeneity, state dependence, and dual taxonomy theories. The main dispute among these contrasting theoretical paradigms centers largely on predicted differences regarding the degree to which criminal propensity is stable/unstable across the life course. This study examines three key questions concerning patterns of continuity and discontinuity across the life courses of serious youthful offenders: (1) how stable are individual differences in the propensity to commit criminal acts across the life course; (2) are there two (or more) discrete groups of offenders with distinct age-crime curves; and (3) is criminal activity of adjacent ages causally related, after controlling for persistent individual differences in the propensity to offend?

Three large data sets of serious youthful offenders released from the California Youth Authority were used in this study. The dependent variable consisted of the number of arrest charges during each age-year. The age ranges considered in this study varied by sample, but overall fell between the ages of 7 and 43. Semiparametric finite mixture Poisson models, as well as parametric random effects and standard negative binomial models were estimated to examine the issues.

The substantive results were identical across all three samples and indicated that: (1) between-group differences were largely unstable across the ages studied; (2) there were more than two discrete groups of offenders found within each sample; and (3) even after accounting for persistent unobserved heterogeneity in the propensity to engage in criminal activities, there was still a significant, positive, and substantively important relationship between having been arrested at the prior age and the frequency of arrest at the current age. The broad substantive implications of these results are that *change matters* in the lives of serious offenders, and even in the lives of the most persistent serious youthful offenders.

ACKNOWLEDGEMENTS

The data collection for this study was made possible by grant # 98-CE-VX-0026 to the California Department of the Youth Authority and the University of California at Davis from the National Institute of Justice. The data collection was facilitated through the cooperation of the Research Department of the California Department of the Youth Authority. The aid of Lee Britton, Rudy Haapanen and Norm Skonovd is gratefully acknowledged. Hernaldo Baltanado, Randy Bonnell, Kim Dochterman, Karla Haber, Kasie Lee, Kathleen Medina, and Hunter Moje worked diligently to code the data employed in this study. We thank Diane Felmlee for her editorial help and guidance on this report. The authors also wish to thank Ken Land, Angela O'Rand, Phil Morgan, and Ken Spenner for their helpful comments on earlier drafts of this report. Finally, the authors wish to emphasize that they are solely responsible for the contents of this manuscript and that the views and opinions expressed in this study are theirs only and do not necessarily represent the views and opinions of the National Institute of Justice, the California Department of the Youth Authority, or the University of California at Davis.

TABLE OF CONTENTS

Abstract i

Acknowledgements iii

List of Tables ix

List of Figures xii

Chapter

1. **Introduction and Overview** 1
   - The "Great Debate" Concerning the Age-Crime Curve 4
   - The Paradox of Persistence 7
   - Explaining The Paradox of Persistence 9
   - Overview of the Dissertation 11
   - Overview of the Findings 14

2. **Theoretical Framework** 17
   - Introduction 17
   - Population Heterogeneity and State Dependence 18
     - Population Heterogeneity 19
     - State Dependence 21
     - Mixed Theories 23
   - Gottfredson and Hirschi's General Theory of Crime 24
     - Crime versus Criminality 26
     - Criminality and Self-Control 27
     - Explaining Continuity & Discontinuity in Crime 29
     - Explaining the Age-Crime Curve 33
     - Explaining the Relationship of Past to Future Criminal Activity 37
   - Sampson and Laub's Age-Graded Social Control Theory 38
     - The Life Course Perspective 38
     - Varying Informal Social Control Over the Life Course 41
     - Explaining Continuity & Discontinuity in Crime 45
     - Explaining the Age-Crime Curve 49
     - Explaining the Relationship of Past to Future Criminal Activity 51
   - Moffitt's Dual Taxonomy Theory 52
     - The Life-Course-Persistent Group 54
     - The Adolescent-Limited Group 58
     - Explaining Continuity & Discontinuity in Crime 64
     - Explaining the Age-Crime Curve 65

Explaining the Relationship of Past to Future Criminal Activity 68
Theoretical Summary 69
Public Policy Implications 77

3. **Literature Review and Hypotheses 83**
Introduction 83
Studies of the Age-Crime Curve 83
Discussion and Hypotheses Related to the Age-Crime Relationship 99
Studies of the Relationship of Past to Future Criminal Activity 106
Discussion and Hypotheses Related to the Past to Future Crime Relationship 127
Contributions to the Discipline 137

4. **The California Youth Authority 145**
Introduction 145
The California Youth Authority 145
The Youth Authority's Origin 147
The American Law Institute's Model Act 148
A Short History of Major CYA Policy and Legislative Changes from 1941-1992 153
The CYA: 1981-1992 160
Rehabilitative and Parole Services Available at CYA Facilities 176
Education 177
Treatment 179
Training 180
Parole Service Programs 183

5. **Data and Methods 189**
Introduction 189
The Three Release Samples 194
Data Sources and Variables 198
Pre-Release Data: Case Characteristics Information 198
Pre-Release Data: Arrest History Information 203
Post-Release Data: Arrest History Information 205
Post-Release Data: Mortality Information 207
Post-Release Data: Adult Incarceration Information 208
Deriving the Analytical Samples 209
Data Limitations 216
Methodology 220

The Dependent Variable 220
Modeling a Count Variable: The Poisson and Negative Binomial Regression Models 222
The Poisson Regression Model 223
The Negative Binomial Regression Model 226
Accounting for Serial Dependence: Persistent Individual Differences 236
The Parametric Random Effects Negative Binomial Model 239
The Semiparametric Mixed Poisson Model 242
Analytical Approach Employed in Chapter 7 250
Analytical Approaches Employed in Chapter 8 251

6. **Descriptive Summary of Data** 255
Introduction 255
Crime, Arrest, And Incarceration Rates in California 256
Criminal Arrest Histories 262
Age at First Arrest 262
Criminal Arrest Patterns 266
Post-Release Recidivism Rates 273
Background Characteristic Variables 280
Variables Related to the Sample Stay 283
Mortality Experiences 286
Adult Incarceration Experiences 289
Summary 292

7. **Age & Crime Among Latent Classes of Serious Youthful Offenders** 294
Introduction 294
Descriptive Summary of the Age-Crime Relationship 301
Selecting the Optimal Number of Latent Classes 304
Method of Estimation & Global/Local Solutions 305
The Optimal Number of Latent Classes in the Three Samples 310
Statistically Testing the Age Invariance Hypothesis 315
Results for the 1981-82 Sample 321
Latent Class Assignment Percentages and Posterior Probabilities 321
Summary Arrest Charge Information 324
Predicted and Observed Trajectories of Arrest 328
The Adult Prison Experiences of the Latent Classes 334
Results for the 1986-87 Sample 338
Latent Class Assignment Percentages and Posterior Probabilities 338

Summary Arrest Charge Information 340
Predicted and Observed Trajectories of Arrest 344
The Adult Prison Experiences of the Latent Classes 350
Results for the 1991-92 Sample 353
Latent Class Assignment Percentages and Posterior Probabilities 353
Summary Arrest Charge Information 356
Predicted and Observed Trajectories of Arrest 360
The Adult Prison Experiences of the Latent Classes 365
Comparison Of Latent Classes Across the Samples 318
Summary of Results 372
Summary of Substantive Results for the 1981-82 Sample 373
Summary of Substantive Results for the 1986-87 Sample 374
Summary of Substantive Results for the 1991-92 Sample 376
Summary of the Comparison of Latent Classes Across the Samples 377
Discussion 378

8. On the Relationship of Past to Future Criminal Activity 390
Introduction 390
Results for the 1981-82 Sample 395
Stage One—The Semiparametric Mixed Poisson Model 395
Stage Two—The Parametric Random Effects Negative Binomial Model 404
Stage Three—Incorporating Latent Class Indicator Variables 407
Stage Four—Latent Class-Specific Models 411
Stage Five—Post-Release Data Only 415
Results for the 1986-87 Sample 422
Stage One—The Semiparametric Mixed Poisson Model 422
Stage Two—The Parametric Random Effects Negative Binomial Model 428
Stage Three—Incorporating Latent Class Indicator Variables 430
Stage Four—Latent Class-Specific Models 432
Stage Five—Post-Release Data Only 435
Results for the 1991-92 Sample 440
Stage One—The Semiparametric Mixed Poisson Model 440
Stage Two—The Parametric Random Effects Negative Binomial Model 445

Stage Three—Incorporating Latent Class Indicator Variables 447
Stage Four—Latent Class-Specific Models 449
Stage Five—Post-Release Data Only 452
Summary of Results 458
Summary of Results for the 1981-82 Sample 458
Summary of Results for the 1986-87 Sample 460
Summary of Results for the 1991-92 Sample 461
Discussion 463

9. **Concluding Remarks** 425
Introduction 425
Theoretical Implications 425

Appendix A **Offenses Seriousness Hierarchy & Categorization of Offense Types** 486

Appendix B **Obtaining Mortality Data from the California Death Statistical Master Files** 489

Appendix C **Percentage of Samples "At Risk" Over the Age Distribution** 496

Appendix D **Cumulative Ages of Onset of First Criminal Arrest** 498

Appendix E **Means of Background Characteristic Variables, by Latent Class** 500

Appendix F **Observed and Predicted Arrest Trajectories, By Latent Class** 504

References 508

LIST OF TABLES

Table 3.1 Summary Information of Prior Studies of the Age-Crime Relationship 84

Table 3.2 Summary Information of Prior Studies of the Relationship Between Prior and Future Criminal Activity 110

Table 4.1. Characteristics of the California Youth Authority and First Commitments, by Year 159

Table 4.2. Parole Consideration Date Guidelines (Months Until Parole), by YOPB Category 164

Table 4.3. Average Months Until Parole at Initial Parole Consideration Date (PCD) Hearing with the YOPB, by YOPB Category and Year of PCD Hearing 168

Table 4.4. Average Length of Stay (in months), by Commitment Offense Type and Year of Release 170

Table 4.5. Assault Rate Against Staff, Assault Rate Against Wards, Total Assault Rate, and Total Disciplinary Infraction Rate, by Year 172

Table 5.1 Constructing the Analytical Samples: Limiting the Release Samples by Sex, M Case Status, and Missing Arrest Data, by Release Sample 210

Table 6.1 Mean and Percentiles of Age at First Criminal Arrest Distribution, by Sample 264

Table 6.2 Sums, Means, and Participation Rates of Arrest Charges, by Sample 267

Table 6.3 Estimates Survival Probability of Remaining "Arrest Free" of A New Criminal Charge as a Function of Years Since Release 275

Table 6.4 Means of Background Characteristic Variables, by Sample 281

Table 6.5 Means of Variables Related to Sample Stay At CYA, by Sample 284

Table 6.6 Frequency and Type of Death, by Sample 287

Table 6.7 Summary of Variables Related to Adult Incarceration Experiences, by Sample....290

Table 7.1 BIC Statistics for the Semiparametric Mixed Poisson Models, by Sample....312

Table 7.2 Parameter Estimates From the 6-Class Semiparametric Mixed Poisson Model, by Sample....314

Table 7.3 Wald Statistics Testing the Equality of the Estimated Parameters Across the Six Latent Classes, by Sample....318

Table 7.4 Summary Descriptions of Latent Class Assignment Percentages and Posterior Assignment Probabilities: 1981-82 Sample....322

Table 7.5 Summary Arrest Charge Information, by Latent Class: 1981-82 Sample....325

Table 7.6 Summary of Variables Related to Adult Incarceration Experiences, by Latent Class: 1981-82 Sample....336

Table 7.7 Summary Descriptions of Latent Class Assignment Percentages And Posterior Assignment Probabilities: 1986-87 Sample....339

Table 7.8 Summary Arrest Charge Information, by Latent Class: 1986-87 Sample....341

Table 7.9 Summary of Variables Related to Adult Incarceration Experiences, by Latent Class: 1986-87 Sample....351

Table 7.10 Summary Descriptions of Latent Class Assignment Percentages and Posterior Assignment Probabilities: 1991-92 Sample....354

Table 7.11 Summary Arrest Charge Information, by Latent Class: 1991-92 Sample....357

Table 7.12 Summary of Variables Related to Adult Incarceration Experiences, by Latent Class: 1991-92 Sample....366

Table 8.1 Investigation of State Dependence Effects With Semiparametric Random Effects Poisson Model: 1981-82 Sample....398

Table 8.2 Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1981-82 Sample....406

Table 8.3 Investigation of State Dependence Effects With NB1 Negative Binomial Models: 1981-82 Sample, Latent Class-Specific Models....413

Table 8.4 Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1981-82 Sample, Post-Release Data Only....418

Table 8.5 Investigation of Subject Characteristic Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1981-82 Sample, Post-Release Data Only....420

Table 8.6 Investigation of State Dependence Effects With Semiparametric Random Effects Poisson Model: 1986-87 Sample....424

Table 8.7 Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1986-87 Sample....429

Table 8.8 Investigation of State Dependence Effects With NB1 Negative Binomial Models: 1986-87 Sample, Latent Class-Specific Models....434

Table 8.9 Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1986-87 Sample, Post-Release Data Only....436

Table 8.10 Investigation of Subject Characteristic Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1986-87 Sample, Post-Release Data Only....439

Table 8.11 Investigation of State Dependence Effects With Semiparametric Random Effects Poisson Models: 1991-92 Sample....442

Table 8.12 Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1991-92 Sample....446

Table 8.13 Investigation of State Dependence Effects With NB1 Negative Binomial Models: 1991-92 Sample, Latent Class-Specific Models....451

Table 8.14 Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1991-92 Sample, Post-Release Data Only 453

Table 8.15 Investigation of Subject Characteristic Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1991-92 Sample, Post-Release Data Only 456

Table A.1 Table of Offense Seriousness Looping Order and Categorization Of Offense Types 487

Table C.1 Percentage of Cases "At Risk" of Arrest at Each Age, by Sample....497

Table E.1 Means of Background Characteristic Variables, by Latent Class: 1981-82 Sample 501

Table E.2 Means of Background Characteristic Variables, by Latent Class: 1986-87 Sample 502

Table E.3 Means of Background Characteristic Variables, by Latent Class: 1991-92 Sample 503

LIST OF FIGURES

Figure 1.1 Violent and Property Aggregate Age-Crime Curves in the United States in 1980, 1994, 2000 2

Figure 2.1 Graphical Representation of the Gottfredson and Hirschi Argument Concerning the Age-Crime Relationship, by Offender Type 72

Figure 2.2 Graphical Representation of the Sampson and Laub Argument Concerning the Age-Crime Relationship, by Offender Type 74

Figure 2.3 Graphical Representation of the Moffitt Argument Concerning the Age-Crime Relationship, by Offender Type 76

Figure 5.1 Graphical Depiction of Routes of Entrance into the California Youth Authority 192

Figure 5.2 Approximating a Continuous Distribution with a Finite Number of "Points of Support" 238

Figure 6.1 California Index Crime Rates from 1960 through 2000, by Crime Type 257

Figure 6.2 Juvenile and Adult Arrest Rates from 1960 through 2000 259

Figure 6.3 California's Adult Incarceration Rate (CDC Prisoners), 1960-2000 261

Figure 6.4 Estimated Survival Probabilities of Remaining Arrest Free as a Function of Years Since Release 278

Figure 7.1 Aggregate Age-Crime Curve, by Sample 302

Figure 7.2 Predicted Trajectories: 6-Class Model of the 1981-82 Sample 329

Figure 7.3 Observed Trajectories: 6-Class Model of the 1981-82 Sample 330

Figure 7.4 Predicted Trajectories: 6-Class Model of the 1986-87 Sample 345

Figure 7.5 Observed Trajectories: 6-Class Model of the 1986-87 Sample 346

Figure 7.6 Predicted Trajectories: 6-Class Model of the 1991-92 Sample 361

Figure 7.7 Observed Trajectories: 6-Class Model of the 1991-92 Sample 362

Figure 7.8 Graphs of Comparable Latent Class Trajectories Across the Three Samples 369

Figure 8.1 Predicted Mean Arrest Charges, by Arrest Status at Prior Age: 1981-82 Sample 402

Figure 8.2 Predicted Mean Arrest Charges, by Arrest Status at Prior Age: 1986-87 Sample 426

Figure 8.3 Predicted Mean Arrest Charges, by Arrest Status at Prior Age: 1991-92 Sample 444

Figure B.1 Graphical Representation of the Mortality Data Retrieval Process ...491

Figure D.1 Cumulative Probabilities of Age at First Criminal Arrest, by Sample.....499

Figure F.1 Observed and Predicted Trajectories of Each Latent Class: 1981-82 Sample 501

Figure F.2 Observed and Predicted Trajectories of Each Latent Class: 1986-87 Sample 502

Figure F.3 Observed and Predicted Trajectories of Each Latent Class: 1991-92 Sample 503

CHAPTER 1

INTRODUCTION AND OVERVIEW

Criminologists generally agree that age is one of the strongest predictors of criminal activity (along with gender) and that a disproportionately large share of offending is committed by those who are in the age cohort 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."

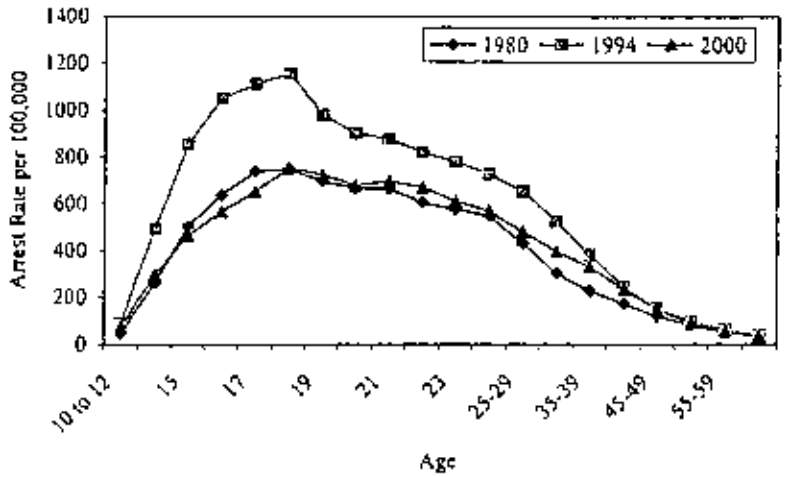
When criminologists speak of the relationship between age and crime, they usually are referring to the "age-crime curve." The aggregate age-crime curve (which is computed by dividing the total number of arrests of individuals of a given age by the total population size of the specific age) indicates: (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, both characterized by 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 curve for both violent and property FBI Index crimes in 1980, 1994, and 2000.[1]

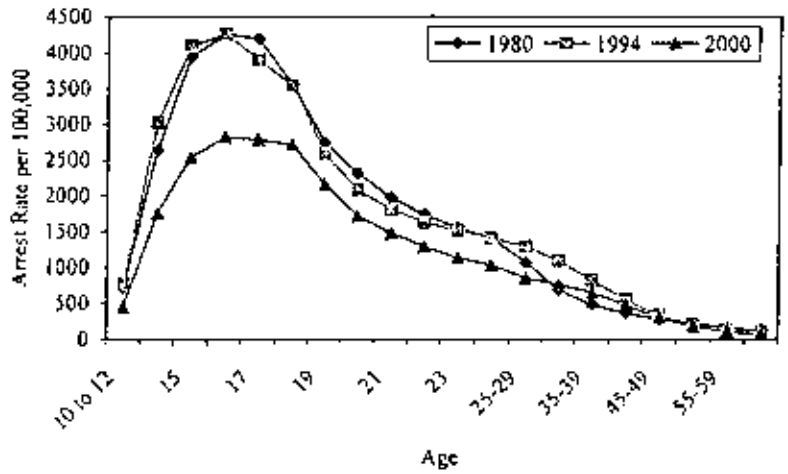
[1] Violent Index crimes include murder and nonnegligent manslaughter, forcible rape, aggravated assault, and robbery. Property Index crimes include burglary, larceny theft, motor vehicle theft, and arson.

Figure 1.1 Violent and Property Aggregate Age-Crime Curves in the United States in 1980, 1994, and 2000

Panel A: Age-Specific Arrest Rates for Violent Index Crime



Panel B: Age-Specific Arrest Rates for Property Index Crime



Source: Adapted from Snyder, H., *Juvenile Arrests 2000*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention, 2002.

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 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 (Hirschi and Gottfredson 1983; Cohen and Land 1987; 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 Laurit:en (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 youthful offenders, examines three key questions related to the relationship between age and crime: (1) how stable are individual differences in the propensity to commit criminal acts across the life course; (2) are there two (or more) discrete groups of offenders with distinct age-crime curves concealed within the aggregate age-crime curve; and (3) is criminal activity of adjacent periods or ages causally related after controlling for persistent individual differences in the propensity to offend?

The "Great Debate" Concerning the Age-Crime Curve

Beginning in the mid 1980s, the field of criminology witnessed what Vold, Bernard, and Snipes (1998: 285) called the "Great Debate" concerning the relationship between age and crime. This debate involved a rather bitter dispute over whether one finds the same relationship between age and crime with individual-level data that is observed when analyzing aggregates. 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 Cohen 1979, 1987; Blumstein et al. 1986, 1988a, 1988b; Farrington 1983, 1986).

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, tend to commit less crime as they age regardless of offense type. Hirschi and Gottfredson further argue that age-specific offense rates increase dramatically from age 10 until age 17, and then continually decrease thereafter. In

addition, Hirschi and Gottfredson (1986) 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 (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 causing the age-crime curve to take its empirical shape. 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 could significantly reduce the crime rate.

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 less 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. They argue that the sharp incline in the early teen years is largely the result of increasing crime participation rates, that the peak ages are determined by participation rates reaching their height during those years, and that the decline in incidence rates is due largely to offenders terminating their criminal careers (i.e., participation rates are declining). 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 even as recently as 1997, 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."

Currently, 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 from purely focusing on whether the relationship between age

and crime is constant among active offenders.[2] Contemporary controversies related to the relationship between age and crime focus on three key questions: (1) how stable or unstable are individual differences in criminal behavior across the 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. In fact, these controversies can all be viewed within the bounds of the longstanding "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. While looking *forward* or *prospectively* in the lives of juvenile delinquents, however, 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: 61) perhaps best

[2] Chapter 3 reviews the evidence from longitudinal studies of criminal offending patterns over time. In short, the evidence supports the fact that crime decreases as a function of age, and refutes the constancy of crime among active offenders position.

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 relationship, and a number of juvenile offenders are able to escape the criminal lifestyle and do not become "career criminals."[3] 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 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 discontinuity in criminal offending patterns must consequentially produce an explanation that is consistent with the observed shape of the age-crime curve and if offending at adjacent ages is causally related after accounting for individual differences in the propensity to offend.

[3] The terms juvenile offender and juvenile delinquent will be used interchangeably throughout this study. While recognizing that there are certain non-criminal behaviors that could cause an individual to be labeled as a delinquent (e.g., status offenses such as running away and incorrigibility), these terms will be used interchangeably herein to reflect behavior that is considered criminal in nature (e.g., theft, assault, burglary, robbery).

Explaining the Paradox of Persistence

Three primary theoretical explanations have been proffered to explain both the paradox of persistence and the shape of the age-crime curve: population heterogeneity, state dependence, and 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 Paternoster 1991; Nagin and Farrington 1992a, 1992b). This latent variable is 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, end their offending earlier, 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 propensity eliminates any relationship between crime at any two points in time

Gottfredson and Hirschi's (1990) self-control theory is 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 it 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]

The dual taxonomy explanations argue that the criminal offender population is comprised of two empirically distinct offender categories, each with its own etiological explanation. In the 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

[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. In its purest form, state dependence theories include labeling theory (Lemert 1972), social learning theory (Akers 1985), and general strain theory (Agnew 1992).

result of desistence by the "adolescent-limited," 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.

Overview of the Study

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

[5] Previous research too often treats this group of serious offenders as a homogenous group (especially in terms of the continuity of their behavioral patterns) because they only comprise a very small fraction (<5%) of general population samples (e.g., birth cohort samples, school samples). They are usually treated as a homogenous group via a binary variable (usually labeled "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 nonoffenders. Sampson (2000) criticizes this static, between-individual research (which produces conclusions indicating high-rate offenders have elevated risks of continuity in offending relative to nonchronic offenders) because it tends to reify the fixed categories and neglects the nature and extent of within-individual change that is occurring with the unfolding of time

(1993). The second chapter also includes a discussion of the public policy implications of these issues. 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 history, policies and practices of the California Youth Authority, the state agency responsible for housing the most seriously delinquent/criminal youthful offenders in the state. We will pay special attention to the procedures and policies in place from 1981-1992, the historical period in which our samples were incarcerated and paroled.

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 1981-82, 1986-87, and 1991-92. The sources of data and variables used in the analyses are described in this chapter. This chapter concludes with a 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 et al. 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, 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 multimethod approach described by Bushway et al. (1999), who recommend using several different statistical models (each with different assumptions) to estimate the relationship between past and subsequent criminal behavior. This is critical because recent research suggests that conclusions of some previous empirical investigations of this issue were possibly method-specific, thereby bringing into question the robustness and reliability of the prior findings. To the degree 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 et al. 1999). Similarly, replicating findings across multiple samples would lend further support to the robustness of any observed effect.

As 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 and directions for future research concerning the topic of continuity and discontinuity in criminal offending patterns.

Overview of the Findings

In general, the results presented in this study lend support to the arguments that (1) there is a significant amount of heterogeneity in the propensity to offend within the serious youthful offender population, and (2) that change is relevant in the lives of serious youthful offenders even after controlling for persistent individual differences in the propensity to engage in criminal activities across the life course.

The results in Chapter 7, which are based on the semiparametric mixed Poisson model of Nagin and Land (1993), indicate that a model with six components in the mixing distribution (or six latent classes) provided the best fit to the data in all three samples. The results indicated significant support for the hypothesis that there are multiple, distinct offender groups on the high-end of the criminal propensity continuum. These findings also provided evidence refuting the claims of Moffitt (1993) that there are only two discrete offender groups concealed within the offender population. Examination of both the observed average total arrest charges and the observed and predicted arrest trajectories of each latent class indicated that there was simply too much heterogeneity in the population (both in terms of the mean rates of offending and the developmental shapes of the arrest trajectories) to be adequately and sufficiently accounted for with only two latent classes. However, the examination of the predicted and observed arrest trajectories in all three samples provided overwhelming support for the presence of an adolescent-limited offender group (consistent with the predictions of Moffitt's theory).

The results presented in Chapter 7 *for all three samples* also send a vigorous signal indicating a lack of support for the age-invariance hypothesis of Gottfredson and Hirschi (1990). The age invariance hypothesis was first statistically rejected using the Wald statistic that tests the restriction of constraining each age parameter to be equivalent across the latent classes. The age invariance hypothesis was then tested substantively by examining the observed and predicted arrest trajectories of the latent classes. The results provided strong evidence of a breakdown in the maintenance of between-group differences across time. In all three of the samples, there was a considerable amount of change in the between-group differences through the late adolescent and adult years studied here.

The results presented in Chapter 8 indicate that even after accounting for the population heterogeneity in the propensity to engage in criminal activities (as measured by arrest data) through both parametric and nonparametric methods, there was still a significant positive relationship between having been arrested at the prior age and the frequency of arrest at the current age. The results also indicated that it was absolutely critical to adequately control for the differences in criminal propensity when estimating the relationship between past and subsequent criminal activity. There was a 50-60% reduction in the magnitude of the state dependence relationship after controlling for persistent individual differences. Models estimated within each latent class failed to uncover significant differential state dependence effects that were stronger in the adolescent-limited group as predicted by the dual taxonomy theory of Moffitt (1993). Overall, the evidence presented in Chapter 8 overwhelmingly favors the *mixed* position that allows for the general importance of both population heterogeneity and state

dependence processes in the explanation of both continuity (stability) and discontinuity (change) in criminal offending patterns across the life course.

Finally, the results presented in Chapter 8 also suggest a significant methodological theme on this topic—a failure to accurately capture the age effects within a sample of data will lead to an overestimation of the estimated state dependence effect. The methodological contribution suggested from the results obtained in this chapter is that it is absolutely critical for researchers to ensure that the age effects are adequately modeled because unaccounted for variation in such effects was found to mask genuine state dependence effects.

CHAPTER 2

THEORETICAL FRAMEWORK

INTRODUCTION

As noted in Chapter 1, this study focuses on three issues central to the continuity and discontinuity of the criminal careers of serious youthful offenders across the age span: (1) the relative stability of criminal propensity over the life course, (2) the degree to which the observed age-crime curve conforms with crime patterns exhibited by multiple heterogeneous groups of offenders with different crime trajectories, and (3) whether there is a relationship between past and subsequent criminal activity after controlling for persistent individual differences among offenders. This chapter focuses on the theoretical framework guiding this study and consists of three main sections. First, the more general population heterogeneity and state dependence explanations will be described in detail. Second, three specific theoretical frameworks and their explanations of these aforementioned issues will be discussed: Gottfredson and Hirschi (1990), Sampson and Laub (1993), and Moffitt (1993). Attention in this section will focus on how each theoretical framework explains continuity and discontinuity of criminal offending patterns, the relationship between age and crime, and the relationship between past and subsequent criminal activity. The third section will discuss the public policy implications of both continuity/discontinuity in crime and the relationship between age and crime.

It is important to note at the outset, that this study does not test the specific causal structures of a particular theory or set of theories, but rather it presents an empirical evaluation of the precise longitudinal implications of three leading criminological

theoretical frameworks noted above. Thus, this study is best viewed as providing evidence either supporting or refuting the direct empirical implications of each of these theoretical frameworks. To date, these implications have remained largely untested among the serious youthful offender population.

POPULATION HETEROGENEITY AND STATE DEPENDENCE

As indicated in Chapter 1, etiological explanations concerning continuity of criminal behavior over time are broadly defined in terms of their basic presumption of either population heterogeneity and/or state dependence processes. Heckman (1981: 150) sharply describes the distinction between the two processes:

> One [explanation] is that individuals who experience the event are altered by their experience in that the constraints, preferences, or prices (or any combination of the three) that govern future outcomes are altered by past events. Such an effect of past occurrences is termed structural state dependence. A second explanation is that individuals differ in some unmeasured propensity to experience the event and this propensity is either stable over time, or if it changes, values of propensity are autocorrelated. Broadly defined, the second explanation is a consequence of population heterogeneity.

Drawing heavily on the "urn schemes" analogies presented in Heckman (1981) and Nagin and Paternoster (1991, 2000), this section explicates the basic principles of both the population heterogeneity and state dependence explanations as they pertain to crime.

To begin, assume that each individual in the population has an urn containing both red and blue balls. The balls represent an individual's propensity to engage in crime and prosocial activities respectively. Over time, individuals pick balls (i.e., event trials)

and replace the balls in the urn (i.e., sample with replacement). Further, allow the drawing of a red ball to represent the event of "committing a crime," while drawing a blue ball represents the probability of engaging in "prosocial activity." The proportion of red balls in an individual's urn represents their criminal propensity. Thus, individuals with greater proportions of red balls in their urns have greater propensities to engage in criminal activities, whereas individuals with greater proportions of blue balls have greater propensities to engage in "conventional or prosocial" behaviors (Nagin and Paternoster 2000: 120).

Population Heterogeneity

Consider first the population heterogeneity urn scheme. According to this perspective, individuals are assigned urns, and the initial constellation of red balls to blue balls *varies across urns in the population*; in other words, there is population heterogeneity with respect to the mix of red and blue balls in individual urns. The critical assumption of the population heterogeneity argument is that for any given individual, once an urn is assigned, the proportion of red and blue balls is considered *fixed across time, stable across time, or time-invariant*. Individuals draw and replace balls over time, but neither red nor blue balls are added to or extracted from a person's urn. Again, individuals with higher percentages of red balls are at greater risk of engaging in criminal activities.[1]

[1] The nature of the processes that cause or generate the initial distribution of red and blue balls in a given urn (i.e., the causes of criminal propensity) and at what point they are considered fixed varies across different population heterogeneity theories, but the basic principles of this theoretical stream of thought are the same.

Given the assumptions of sampling with replacement and the fixed nature of red and blue balls in any given urn, the odds of drawing a red ball (i.e., committing a crime) or a blue ball (i.e., engaging in prosocial activities) *never* change for each individual across the life span. Accordingly, some individuals in the population simply have a greater chance of drawing a red ball (because they have more red balls in their urn) and, even more importantly, these individuals are consistently more likely to engage in criminal behaviors relative to those who have more blue balls in their urn because the distribution of red to blue balls never changes. In other words, each draw from the urn is statistically independent of the prior draws in the sense that drawing a red ball does not increase (or decrease) the odds of drawing another red ball at the next draw. In retrospect, however, the knowledge of an individual's past experience of drawing red balls will certainly be highly predictive of the odds of a future drawing of a red ball. For example, an individual who has never drawn a red ball will be unlikely to draw a red ball in the future. Conversely, an individual who has only picked red balls in the past is highly likely to continue picking red balls in future trials. Why is this so? According to the population heterogeneity perspective, the correlation between past and future draws is simply determined by the initial mix of red and blue balls in one's urn.

As Nagin and Paternoster note (2000: 121), "the predictive power of past events is entirely due to the *initial distribution* of red and blue balls in the urn" (emphasis in original). For example, any observed correlation between past and subsequent criminal activity is entirely due to the initial distribution of red balls in the urn. Since the probability of picking a red (or blue) ball is constant over time (i.e., from trial to trial), continuity in behavioral patterns (criminal or prosocial), is simply a consequence of the

initial propensity to engage in those behaviors (i.e., the initial odds of drawing a red or blue ball). Thus, the occurrence of completing high school, going to college, involving oneself in a stable marriage, and obtaining a job are all seen to arise as a consequence of the initial distribution of the propensity to engage in such conventional acts. Individuals experience these events as a consequence of the initial number of blue balls in their urns. Similar to the lack of causation between prior and subsequent criminal activity, there are no causal links between the occurrence of prosocial activities and future criminal behavior. Again, the correlations between the number of red and blue balls drawn in the past and the odds of drawing a red or blue ball in the future are entirely determined by the initial distribution of the red and blue balls in one's urn. The correlation is not causal, but rather it simply (and spuriously) reflects the initial distribution in the individual urns.

Under the assumption that it is not possible to see directly inside the urn to count the exact number of red and blue balls, past counts of red balls drawn can be used as an indicator of a given individual's latent criminal propensity. Naturally the greater the number of trials observed (i.e., the longer the length of follow-up period in a study of criminal behaviors), the more accurately one could measure the latent propensity variable.

State Dependence

Using the same urn analogy, the pure form of state dependence differs from the population heterogeneity explanation on two key assumptions. First, all individuals are initially assigned identical urns with the exact same number of red and blue balls. This contrasts sharply with the population heterogeneity explanation because all individuals

are viewed there as having equal odds of initially selecting a red or blue ball (i.e., all individuals have equal criminal propensities in this model).

Second, while individuals still sample with replacement, the number and mix of red and blue balls is malleable over time. The critical assumption of the state dependence perspective is that the selection of a given ball results in the addition of one or more balls of the same color drawn in the trial to the individual's urn. Thus, if one selects a red ball, that drawn red ball is replaced and additional red balls are deposited in the individual's urn. In other words, the commission of a criminal act (i.e., the selection of the red ball) is argued to causally increase the odds of future criminal acts (i.e., additional red balls are added to increase the proportion of red balls in the individual's urn). The same process is assumed to occur for the selection of a blue ball; engaging in a conventional or prosocial activity is argued to increase the odds of future prosocial activities.

Thus, the state dependence position views continuity in behavior as resulting from the fact that after the event of picking a ball of a given color, the odds of picking that color increase in the future because of the additional balls added to the urn. In other words, the proportion of red and blue balls is considered to vary over time and to be causally related to past events. Regarding the relationship between past and subsequent criminal activities, the observed correlation is argued to be genuinely causal in nature rather than spuriously due to the initial distribution of red and blue balls as in the population heterogeneity explanation. Thus, one can see why population heterogeneity theories are often referred to as *static* theories while theories assuming a state dependence process are often referred to as *dynamic* theories (Paternoster et al. 1997). The key proposition of state dependence theories then is that events have consequences, and that

these consequences can either increase or decrease the likelihood of future criminal behavior.

> The state dependence position adopts the view that: (1) criminal behaviors may subsequently open up new opportunities for other criminal activities while closing off opportunities for noncrime, and (2) some noncriminal behaviors may subsequently open up opportunities for other noncriminal behaviors while closing off criminal opportunities (Nagin and Paternoster 2000: 125).

As a dynamic perspective, state dependence theories directly imply that even if in the past one has engaged in criminal activities (and added more red balls to one's urn), engaging in prosocial activities can decrease the probability of future criminal activity at any point because more blue balls will be added to the person's urn. In other words, criminal propensity can be significantly altered over the life course by continued involvement or investment in prosocial activities (Nagin and Paternoster 1993, 1994).

Mixed Theories

While in their pure forms population heterogeneity and state dependence explanations are diametrically opposed explanations, they are not mutually exclusive processes (Nagin and Paternoster 2000; Sampson and Laub 1997). That is, theories can both allow for population-level heterogeneity in the initial distribution of criminal propensity, while also allowing for consequences to result from engaging in criminal activity. As Sampson and Laub (1997: 155) state, "to assume that individual differences influence the choices one makes in life (which they certainly do), does not mean that social mechanisms emerging from those choices can then have no causal significance." Indeed there is a growing consensus in the field of criminology that persistent individual

differences must be incorporated into any valid theoretical explanation of criminal behavior since it has become obvious that "there are persistent differences across individuals in the rates of offending over time" (Land and Nagin 1996: 164). Whether those persistent differences are the "be all and end all" of explaining crime and exactly how stable they are over time is a fundamental debate in criminology (Paternoster, Brame, and Farrington 2001).

In the next three sections, the theories of Gottfredson and Hirschi, Sampson and Laub, and Moffitt representing the population heterogeneity (Gottfredson and Hirschi) and mixed models (Sampson and Laub, Moffitt) are described. [2] A key distinction between each of these theories is in their different explanations of the stability of antisocial tendencies over the life course. The question of stability has direct implications for explaining the shape of the age-crime curve. Thus, particular attention is focused on each theory's explanation of continuity in crime and its corresponding explanation of the age-crime relationship.

GOTTFREDSON AND HIRSCHI'S GENERAL THEORY OF CRIME

In their book *A General Theory of Crime* (1990), Gottfredson and Hirschi explicate their population heterogeneity theory centered-around the notion of self-control.[3] This theory has profound implications for sociological theories of crime

[2] Even though both Sampson and Laub's (1993) age-graded informal social control theory and Moffitt's dual taxonomy theory are conceptually distinct theories, both theories are examples of integrating elements of both population heterogeneity and state dependence propositions into their etiological explanations of criminal behavior.

[3] Wilson and Herrnstein (1985) have proposed another well-known population heterogeneity theory. The causes of criminal propensity vary somewhat between their theory and the theory of Gottfredson and Hirschi. Gottfredson and Hirschi place the root causes of criminal propensity in the early child rearing

because it challenges the basic fundamental premises of most sociological theories (Cohen and Vila 1996; Nagin and Paternoster 1994), and thus their theory is highly deserving of both empirical testing and critical evaluation.

To begin, Gottfredson and Hirschi say that their theory is "meant to explain all crimes, at all times, and for that matter, many forms of [risk taking] behavior that are not sanctioned by the state" (1990: 117), which is why they refer to their theory as a *general* theory of crime. Gottfredson and Hirschi place the concept of self-control as the centerpiece around which nearly every "fact" of crime can be organized and explained, including continuity in crime, the age-crime relationship, the gap between male and female involvement in criminal activities, the disproportionate involvement of minorities, the role of peer groups, why prosocial activities are negatively correlated with criminal activity, and why criminal offenders tend to engage in a constellation of noncriminal yet similarly risk-taking behaviors that are "analogous to crime" (e.g., alcoholism, drug abuse, smoking, excessive speeding in an automobile, automobile accidents, promiscuous and unprotected sexual activity).

---

practices of the caretaker. While agreeing that child rearing practices are very important, Wilson and Herrnstein theory's, however, is more strongly rooted in the biological and constitutional determinants of criminal propensity, including a genetic predisposition towards criminal propensity. According to their theory, biological and constitutional factors, along with child-rearing techniques, induce population variation in the degree to which people are more or less present-oriented. Individuals who are more present- rather than future-oriented will be more inclined to engage in criminal activity because the rewards of crime come immediately, whereas the costs are generally deferred to some point in the future. Importantly, Wilson and Herrnstein argue that the inability to defer gratification is established early in life and is a stable personality trait across time that is impervious to change. Thus, their theory makes similar longitudinal implications about crime continuity as the theory of Gottfredson and Hirschi.

Crime versus Criminality

One of Gottfredson and Hirschi's primary theoretical contributions is their argument for distinguishing between crime and criminality as a necessity for understanding the etiology of crime. They argue that the failure of positivistic etiological explanations to make this distinction renders most theories of crime seriously flawed in the conceptualization of their dependent variable (Gottfredson and Hirschi 1990: 144). In short, For Gottfredson and Hirschi, crime refers to the behavioral acts that people engage in, whereas criminality refers to the individual's propensity to engage in crime.

To be consistent with the "characteristics of ordinary crimes", Gottfredson and Hirschi define crime as "acts of force or fraud undertaken in the pursuit of self-interest" (1990: 15-16). According to them, crimes are simple behavioral acts that: (1) involve immediate gratification and satisfy ordinary and universal desires; (2) provide few long-term benefits to the actor and cause pain and suffering to the victim; and (3) are exciting, risk-taking behaviors that can be committed by every individual in society without specialized knowledge, training, or prior learning.

While most criminological theories try to explain the "causes of crime", they fail to clearly conceptualize their dependent variable-- crime. By distinguishing between crime and criminality, Gottfredson and Hirschi remove the confounding preoccupation with the "acts of crime" from the real theoretical question of criminality—explaining the propensity of some individuals to engage in crime. This focus on the actor clearly sets the stage for the central mechanism they employ to explain criminality—self-control.[4]

[4] It is important to note that this definition of crime is independent of the formal criminal laws of the state, but encompasses nearly all activities that would be prohibited by the state (Tittle 1995)

Criminality and Self-Control

To describe the type of individual most likely to repeatedly engage in behaviors fitting their definition of crime, Gottfredson and Hirschi offer a *control theory*; they posit (like all control theories) that people would normally be free to commit crimes in the absence of some "controlling" force restraining or preventing them from satisfying their immediate desires (Vold, Bernard, and Snipes 1998).

According to Gottfredson and Hirschi (1990: 90-91), individuals who repeatedly commit these kinds of acts "will tend to be impulsive, insensitive, physical (as opposed to mental), risk-taking, short-sighted, and nonverbal." Importantly, they argue that these traits are positively correlated and tend to coalesce within certain individuals and to *persist* throughout the life course. The authors note that the above characteristics are the precise traits of individuals with low levels of self-control. Thus, their theory posits that the inclination to commit criminal acts in the pursuit of self-interest is a function of an individual's lack of self-control.

Stated succinctly, individuals are assumed to vary in their propensity to use force and fraud as a means of fulfilling their own self-interests and/or obtaining resources; this varying propensity is what Gottfredson and Hirschi refer to as *criminality*. People have varying degrees of criminality because there is a population variation in the level of self-control. People with high-levels of self-control have low-levels of criminality, whereas individuals with low levels of self-control have high levels of criminality.

Again, the authors also argue that their theory explains not only criminal acts, but other "analogous" behaviors as well (e.g., alcoholism, drug abuse, automobile accidents).

Individuals with low self-control tend to engage in a wide variety of risky acts consistent with the definition of "crime." Further, individuals with low self-control also have difficulty, for example, obtaining and keeping employment, maintaining friendships with spouses and others, completing educational endeavors, becoming effective parents, staying healthy, and keeping long-term financial commitments. In other words, low self-control does not enhance the quality (or longevity) of life and severely restricts the potential to have or build positive social relationships. As Gottfredson and Hirschi (1990: 96) argue,

> Social life is not enhanced by low self-control and its consequences. On the contrary, exhibition of these tendencies undermines group relations and the ability to achieve collective ends. These facts explicitly deny that a tendency to commit crime is a product of socialization, culture or positive learning of any sort.

If the propensity to engage in crime is a consequence of low self-control, the obvious question is: what are the causes of these varying levels of self-control? According to the Gottfredson and Hirschi, levels of self-control arise largely from family teaching and child rearing practices. Effective child rearing includes the following three components: (1) the parents must adequately monitor the child's behavior; (2) deviant behavior must be recognized when it occurs; and (3) deviant behavior must be consistently punished when it is recognized (Gottfredson and Hirschi 1990: 97). These three external controls on the child's behavior when employed consistently by effective parents eventually become internalized in the child through the process these authors term "socialization." Effective socialization develops the abilities to delay gratification, empathize with others, and to sacrifice personal needs for the well-being of others. In the

absence of effective parenting/socialization, individuals will not develop the self-restraint (i.e., high-levels of self-control) necessary to resist pursuing one's self-interest in the easiest, quickest manner possible: through crime; or as Gottfredson and Hirschi (1990: 88) pointedly state, "money without work, sex without courtship, revenge without court delays."

Considering both the concepts of crime and criminality together, Gottfredson and Hirschi claim to have produced an internally consistent argument by making their conception of crime congruent with their conception of criminality: "people who commit crimes are assumed to possess traits that reflect the nature of those acts" (Barlow (1991: 233).

However, this internally consistent result has been criticized as tautalogical because their conception of low self-control is defined by the very behaviors assumed to be indicators of low self-control. For example, Akers (1991) argues that until measures of self-control are operationalized independent of the behaviors said to reflect low self-control, the theory will remain tautological. In their defense of this criticism, Gottfredson and Hirschi (1993: 52-53) retort, "the charge of tautology is a compliment; an assertion that we followed the path of logic in producing an internally consistent result" and that nontautological theories will produce definitions of crime and criminality that are "independent of one another."

Explaining Continuity & Discontinuity in Crime

A crucial proposition of Gottfredson and Hirschi is that criminal propensity is set very early in life (generally by around the age 8), and that the degree of self-control

instilled at that point will remain stable throughout the remainder of the life course. From their perspective, the failure of socialization processes to instill adequate levels of self-control is extremely difficult to overcome in later life, and, similarly, high-levels of self-control once attained cannot be easily unraveled. Again, we emphasize that although the authors clearly favor the position that it is easier to instill higher-levels of self-control than it is to reduce the self-control already instilled, they ultimately contend that criminality is largely determined and intractable by around the age of eight (1990: 106). They do, however, admit that an individual's level of self-control is somewhat flexible over time. They attempt to reconcile the discrepancy between the notion of self-control (criminal propensity) as a time-stable trait on the one hand, and allowing for the possibility of it changing over time on the other hand, in two ways. First, Gottfredson and Hirschi assert that changing from lower- to higher-levels of self-control is perhaps possible, but extremely unlikely because the preexisting low level of self-control and all of its consequences essentially overwhelm the odds of such change. Second, Gottfredson and Hirschi (1990: 107) argue that while "socialization may continue to occur throughout life" for everyone in the population, the relative position one occupies on the self-control distribution scale in the population continues to remain stable throughout life. In other words, they assume that socialization may continue throughout the life course, "the rate at which socialization continues to occur is approximately the same for everyone" (Nagin and Paternoster 2000: 122).

In sum, the relative positions of individuals with respect to the distribution of self-control present in the population at approximately the age 8 will be equal to the relative distribution of criminality present in the population at age 20, age 30, age 40, and so on.

In the words of Gottfredson and Hirschi (1990: 107), "differences between people in the likelihood that they will commit criminal acts persists over time." Barlow (1991: 235) refers to this as the "stability postulate."

It is important to point out that Gottfredson and Hirschi's argument of time-invariant individual differences in criminal propensity after age 8 essentially discredits all modern positivistic theories of crime from the disciplines of psychology, economics, and sociology.[5] If that assumption is true, then experiences in later childhood, adolescence, and adulthood are essentially irrelevant to crime causation, and an individual's educational, economic, social, and psychological experiences cannot have an enduring impact on criminal involvement (Nagin and Paternoster 1994). Moreover, Gottfredson and Hirschi argue that all of those experiences are, in fact, determined by initial levels of self-control and thus none will have any effect on criminal behavior after one controls for this initial level of self-control. Thus they argue that individuals *self-select* such conventional experiences, meaning that individuals will select or choose educational, economic, and social experiences that are entirely consistent with their level of the latent trait variable, self-control (Benson 2002).

[5] It is worth noting that Gottfredson and Hirschi's conceptualization of economic theories of crime (e.g., Beccaria 1963; Becker 1968) as positivistic is at serious odds with the main theoretical propositions of positivistic criminology (see e.g., Jeffrey 1972; Cohen and Land 1987; Vold et al. 1998), and stands in stark contrast with their earlier descriptions of economic theories of crime as distinct from positivistic theories (Gottfredson and Hirschi 1987b). The key propositions of positivistic criminology are that forces beyond the control of the individual determine criminal behavior, that individuals behave as they have been *determined* to do so, that criminals are different (sociologically, psychologically, biologically) from non-criminals, and a *rejection of individual free will*. Broadly defined, economic theories of crime such as the classical and neoclassical models envision individual behavior as determined by free will, that individuals are individually responsible for their own behavior because they freely *choose from available options* , and that criminals are normal rational individuals (not sociologically, psychologically, or biologically different than non-criminals (except in their choice to engage in criminal behavioral options).

For example, the authors argue that having delinquent peers is spuriously associated with criminal activity because "birds of a feather flock together." Delinquent peers are impulsive, reckless youth who are fun to be around because they are adventuresome, thrill-seeking, and present-oriented. As Nagin and Paternoster (1993: 490, 1994) argue, individuals with low self-control have incredibly high "discount rates" whereby such individuals "place less value on future consumption, [so] they are unlikely to invest in a line of activity that sacrifices immediacy for future gratification." Thus, delinquent peer groups form consisting of individuals with low self-control who will take advantage of available risky opportunities, opportunities that are facilitated through peer group interaction. However, due to the nature of low self-control, these groups have short life spans because individuals with low self-control have qualities (e.g., impulsivity, self-centeredness, untrustworthiness) that prevent the lasting formation of any stable relationship among such groups. Low self-control brings them together (not "status deprivation" or "anomie" or "social disorganization"); low self-control will ultimately break them apart. This view of delinquent peer groups stands in stark contrast to the etiological significance of gang maintenance in cultural deviance/social learning and strain theories that envision subcultural peer pressure/social reinforcement as a major cause of criminal activity (e.g., Sutherland and Cressey 1978; Akers 1985).

Similarly, Gottfredson and Hirschi argue that because of such concomitant characteristics (e.g., impulsivity) individuals with low self-control are uncomfortable in structured environments and do not perform well in school or traditional jobs that involve rote tasks. As a result, individuals with low self-control tend to do poorly in school and work environments. They often leave school (before graduating) are unable to retain

employment and search of alternative environments that do not advocate following rules, or being punctual or orderly, attentive, and quiet for extended periods of time. Parallel arguments are made by Gottfredson and Hirschi for the correlation between criminal activity and marriage stability, and parental investment. The correlations between crime and those events/experiences are entirely spurious and due to a lack of self-control itself. Thus, from Gottfredson and Hirschi's perspective, individuals with low-levels of self-control *self-select* or sort themselves over the life course into life experiences and choices (crime, nonmarriage or bad marriages, ineffective parenting, high unemployment, low educational attainment, frequent drug and alcohol abuse) entirely consistent with their deficient level of this underlying characteristic or trait (Nagin and Paternoster 2001).

Explaining the Age-Crime Curve

If the propensity to engage in criminal activity is constant throughout life, how can Gottfredson and Hirschi explain the shape of the age-crime curve, especially the apparent sharp decline in crime after late adolescence? To get around such seeming incongruity, they draw on their initial distinction of crime and criminality and argue that criminality only predisposes people to engage in crime in the face of available criminal *opportunities*. In the absence of opportunities, criminal propensity is just that, criminal propensity. Individuals with high degrees of self-control will rarely commit crimes, even in the presence of opportunities, whereas individuals with low self-control will use force or fraud to pursue their own self-interest at a relatively high rate in the face of such opportunities to do so. The implication of their argument is that because of age roles,

structural factors and perhaps even biological processes, the opportunities to commit criminal acts tend to decline with age.

Thus, since criminal propensity remains fixed across time, the authors clearly rely on a host of different social forces that lead to a "reduction in exposure to criminal opportunities, that, on average, decline as activity patterns change with age," to explain the age-crime relationship (Cohen and Vila 1996: 131). It should be noted, however, that Gottfredson and Hirschi are not as clear as they could be on this topic (Tittle 1995; Cohen and Vila 1996).

For the purposes of this study, though, Gottfredson and Hirschi's critical argument concerns "age invariance" and its longitudinal implications. Their age invariance argument, originally presented in their *American Journal of Sociology* article "Age, Crime, and Social Explanation" (Hirschi and Gottfredson 1983), posits that the relationship between age and crime is "inherent, invariant, and inexplicable" (Tittle and Grasmick 1998). No matter whether one uses self-report offense data or official police or court records of arrest, the data suggest that all people, everywhere, and within any historical period, tend to commit less crime as they age. The authors argue that if the relationship between age and crime is invariant and *all* individuals commit less crime as they age, then age is actually irrelevant to the study of crime and no sociological, psychological, or economic variables that covary with age (e.g., employment, marriage) can explain this "age effect" (Tittle and Grasmick 1998).

Because this is such a controversial argument with far reaching implications for both the explanation and the proper methods for the study of crime, we consider their argument in further detail here. The authors have made it clear in several expositions of

the invariance argument (Gottfredson and Hirschi 1987, 1990; Hirschi and Gottfredson 1983, 1985, 1986, 1988) that they believe the shape of the age-crime curve is relatively robust across persons, groups, cultures, and periods. All sources of data suggest that individuals will have their greatest involvement in criminal activity during the late adolescent years of life and offending will decline thereafter. The implication of this argument is that even individuals with vastly different life circumstances and experiences will have similarly shaped age-crime curves across the life course (Greenberg 1985). However, Greenberg (1985) notes that if the age-crime curve results from the effects of social processes that develop with age, then those processes should affect "different groups differently, breaking the uniformity of the relationship between age and crime across groups." Gottfredson and Hirschi argue, however, that crime declines regardless of whether individuals experience such events as employment, completion of schooling, and marriage, an argument that directly counters the explanations of life course researchers such as Sampson and Laub (1993) and Marxist criminologists such as Greenberg (1985).

Again, we reiterate that the key implication of Gottfredson and Hirschi's invariance argument that we are concerned with here is that the *differences between individuals persist over time.* Both the relative differences of criminal propensity (criminality) and relative group differences in criminal offending (crime) should endure throughout life. Group differences in criminal offending histories reflect "no more than group variation in the propensity to commit offenses at any point in the life course" (Shavitt and Rattner 1988: 1459). Thus, the only explanation needed is why some individuals/groups have higher rates of involvement in crime at any point in time than do

other individuals or groups. For Gottfredson and Hirschi, of course, this is due to the different levels of self-control distributed throughout the population. Tittle and Grasmick (1998: 314) provide an excellent summary of the invariance argument of Gottfredson and Hirschi on this point:

> Variations in criminal behavior between those with different degrees of self-control at any age will be similar to such differences at any other age, even though the absolute amount of crime by everybody changes over the life cycle in conformity with the inverted-J curve.

The implications of the invariance argument are profound and far-reaching. If the relationship between age and crime is invariant and between-group differences that exist at one point in the age-crime curve continue to exist at any other point in the age-crime curve, then only a single time point is necessary to measure the criminality of any group. To quote Gottfredson and Hirschi (1987: 592) "if there is continuity over the life course in criminal activity (or its absence), it is unnecessary to follow people over time." Following individuals across time merely provides redundant information (available at any point in cross-section) at a hefty price because longitudinal research is vastly more expensive to conduct than cross-sectional research (Gottfredson and Hirschi 1987a).[6]

If Gottfredson and Hirschi are wrong, however, and there are different criminal offending trajectories in the population that do not follow the overall aggregate age-crime curve, then it is absolutely necessary for researchers to follow individuals over time to determine not only the actual empirical shape of their crime trajectories, but also if any events or experiences help explain the different trajectory shapes. If crime is a social

[6] Sampson (1992: 546) critiques the Gottfredson and Hirschi's argument that longitudinal data provide no empirical benefits to the study of crime and wastes research money because "such data are necessary to verify the core assertions of their theory regarding stability and the lack of change across the life course."

event that takes on different meanings across the life course (Greenberg 1985; Hagan and Palloni 1988), then it is necessary to study trajectories of criminal offending as dynamic processes that unfold over time, with a particular emphasis on whether trajectories of crime are linked or interrelated with trajectories in other social, economic, psychological, and perhaps even biological domains of life.

Explaining the Relationship of Past to Subsequent Criminal Activity

This discussion of Gottfredson and Hirschi's self-control theory concludes by describing their position on the relationship of past to subsequent criminal activity. Recall that their self-control theory uses a population heterogeneity argument in which Gottfredson and Hirschi argue that the correlation between past and subsequent criminal activity is merely a spurious correlation due to an unmeasured (omitted) variable—level of criminal propensity. It is a classic "variant of the familiar 'omitted variable' bias argument" (Nagin and Paternoster 1991: 166). In their view, individuals with high criminal propensity tend to commit crimes very frequently, including at adjacent ages and/or time periods, while individuals with very low criminal propensity rarely if ever commit criminal offenses. Thus, there will naturally be a high correlation between criminal offending measured at two different time points. The high correlation, however, is argued to arise from the missing variable denoting the level of criminal propensity. If their argument is correct, then including a variable to measure criminal propensity in an equation should eliminate the significant relationship between past and subsequent criminal offending. According to Gottfredson and Hirschi (1987: 594), "subsequent delinquency cannot be predicted among groups homogeneous on current delinquency."

Their assumption that criminal propensity is time-invariant is absolutely critical to their argument, for "if they were not enduring, the [population heterogeneity] theories could not explain the positive association of past to subsequent criminality" (Nagin and Farrington 1992: 237).

SAMPSON AND LAUB'S AGE-GRADED SOCIAL CONTROL THEORY

In contrast to Gottfredson and Hirschi, Sampson and Laub (1993) present an age-graded theory of informal social control that focuses on the *changing/malleable* strength of social bonds over the life course. Their theory draws heavily on the main principles of the life course perspective in sociology (Elder 1985; Riley 1986). Before proceeding to a discussion of their theory, we first present a brief theoretical backdrop describing the life course perspective.

The Life Course Perspective

The life course perspective is both a conceptual and a theoretical perspective (Elder 1992). As a theoretical perspective, the life course "is a theoretical orientation for the study of human development that incorporates temporal, contextual, and processual distinctions" (Elder 1996: 1131). As a concept, the life course refers to "the interdependence of age-graded trajectories, such as work or family, that are subject to changing conditions in the larger world, and to short-term transitions, ranging from birth to school entry to retirement" (Elder 1996: 1121). The life course perspective envisions aging and development as a process that continues throughout life (Riley 1986).

Two key theoretical concepts of the life course perspective are *trajectories* and *transitions*. A trajectory is a longitudinal series or sequence of linked states within a major domain of life (e.g., social, psychological, or biological states) (Elder 1985). In essence, a trajectory is a line of development or pathway over the life span (Sampson and Laub 1990). For example, individuals have educational trajectories, marital trajectories, physical and mental health trajectories, criminal offending trajectories, and employment trajectories just to name a few. Trajectories are long-term patterns of behavior that often exhibit both change and stability depending on whether they are interrupted by transitions.

Transitions are life events that represent discrete changes of state; they evolve over shorter periods of time, and are embedded in trajectories (Elder 1985; Sampson and Laub 1990, 1992, 1993). Some examples of transitions include graduating from high school, getting married, or obtaining a job. Some transitions act as "turning points" because they serve to redirect or change the course of the trajectory (Elder 1985; Sampson and Laub 1993). The long-term view of trajectories implies a strong connection between childhood and adolescence, and between adolescence and adulthood, but the short-term view implies that trajectories can be modified by transitions and even *redirected* by turning points (Laub et al. 1995).

One of the central premises of the life course perspective is that trajectories in different domains of life tend to be interlocked or interconnected because changes or transitions in one domain of life are often associated with changes in other domains. In other words, trajectories can have reciprocal effects on one another (Elder 1985). Indeed, it is this interlocking nature of trajectories that allows for change in one's life course.

Other key premises or themes of the life course perspective include the idea that aging and development cannot be separated from the historical time and place in which it occurs, also known as the principle of contextualism (Dannefer 1984; Elder and O'Rand 1995), the timing of events or the age at which the events occur is crucial for determining the effects of those events on individuals (the life stage principle), and that our lives are linked or embedded in the lives of individuals around us (the linked lives principle) (Elder 1985, 1996). Clearly, those who study the life course consider the acquisition of longitudinal data as imperative to any research design.

In the criminological literature, the life course perspective is considered one branch of what has become known as "developmental criminology" (Vold et al. 1998). The term developmental criminology refers to "within-individual changes in offending" and a major interest of this theoretical paradigm is in the documentation and explanation of longitudinally dynamic patterns of offending from childhood through adulthood (LeBlanc and Loeber 1998: 117). According to Loeber and Stouthamer-Loeber (1996), there are three main goals of developmental criminology: (1) describing within-individual changes in offending patterns over the life course; (2) developing etiological explanations of the longitudinal patterns of offending; and (3) examining the impact of transitions on patterns of offending. On the basis of these three goals, Sampson and Laub's theory definitely qualifies as a developmental theory, for it focuses precisely on these three goals.

In their book, "*Crime in the Making: Pathways and Turning Points Through Life,*" Sampson and Laub (1993) first proposed their age-graded theory of informal social control. The authors embarked on this theoretical exposition to move criminologists past

their preoccupation with adolescence by demonstrating the importance of explaining variation in criminal behavior over the entire life course. As Sampson and Laub note, they were interested in bringing "both childhood and adulthood back into the criminological picture" (1993: 7) because sociological criminology "has not come to grips with the link between early childhood behaviors and later adult outcomes" (Sampson and Laub 1990: 609). It was of special concern to Sampson and Laub to confront and reconcile the "paradox of persistence" phenomenon discussed in Chapter 1.

Varying Informal Social Control Over the Life Course

Sampson and Laub (similar to Gottfredson and Hirschi) posit a control theory that assumes people normally will often commit crimes in the absence of some "controlling" force that restrains or prevents them from engaging in these acts to satisfy their desires. It is the source or locus of the "controlling" force that is the quintessential difference between the two theories. To Gottfredson and Hirschi, the locus of the constraining force is a time-invariant internalized force (self-control) that is fixed after early childhood, whereas for Sampson and Laub the constraining forces (informal social controls) dynamically varies across the life course. Sampson and Laub argue that crime is more likely to occur when an individual's bond to society is weak or broken. Ironically, their theory draws heavily on the notion of social ties developed by Hirschi (1969) in his "social bonding" theory. According to Hirschi (1969: 16), "delinquent acts result when an individual's bond to society is weak or broken." The bond is comprised of various attachments, commitments, involvements and beliefs that when present constrain the individual from attempting to satisfy desires, wants, and needs through illegal means.

Although Hirschi's theory was originally constructed in static terms, Sampson and Laub provide a dynamic interpretation of Hirschi's theory that allows the strength of social bonds to vary over time (Sampson and Laub 1997).[7] This approach allows Sampson and Laub to explain the variation in individual patterns of crime across the life course they observe among their longitudinal data. Such a dynamic conceptualization of changing social bonds over individual lifetimes fits perfectly with the life course perspective.

Utilizing the life course perspective then, Sampson and Laub differentiate the life course by age or life stages and argue that the critical institutions of formal and (especially) informal social control vary across these stages. To Sampson and Laub, the key explanation of differential crime patterns across the life course is the varying amount of informal social control present in childhood, adolescence, and adulthood. Over the life course, the key institutions responsible for varying levels of social control are:

- In childhood and adolescence they are the family, school, and peer groups;
- In young adulthood they are higher education, vocational training, work, and marriage;
- In subsequent adulthood they are work, marriage, parenthood, military service, and investment in the community (Sampson and Laub 1990).

More importantly, Sampson and Laub emphasize the "role of age-graded informal social control as reflected in the structure of interpersonal bonds linking members of

[7] Recall that Hirschi (1969) argued there are four elements to the social bond: (1) *attachment* to others; (2) *commitment* to conventional activities such as school and work which causes individuals to have a stake in conformity; (3) *involvement* in conventional activities, (4) and *belief* in general conventional values, norms, and laws. Thornberry (1989: 876) posits a dynamic "interactional" theory by integrating the principles of Hirschi's (1969) social bond theory with social learning theory (Akers 1985) whereby "delinquency eventually becomes its own indirect cause precisely because of its ability to weakened further the person's bond to family, school, and conventional beliefs."

society to one another and to wider social institutions (e.g., work, family, school)" (Laub et al. 1995: 93). Like Durkheim (1897 1951), Sampson and Laub define social control as the ability of a social group to regulate the behaviors of its members according to its accepted norms and values. Their crucial argument is that the most important sources of social control are actually the informal social bonds that *emerge out of or from* role relationships established for purposes other than social control. Thus, it is not the variability of age-graded social institutions themselves that serve to induce conformity, but rather it is the informal interpersonal bonds between people that serve to link individuals to those institutions; informal social control is not maintained merely by having a teacher or parent present. Conformity to norms is most likely to result when the quality of that relationship between the child and the care giver is high.

The theory of Sampson and Laub thus highlights the quality of interpersonal relations between individuals (e.g., parent-child, student-teacher, husband-wife) as a form of *social investment* or *social capital* (Coleman 1988, 1990), which is created when relationships of interdependence serve to facilitate action and provide social and psychological resources for those individuals to utilize (Laub et al. 1995). There are two critical points associated with the "social investment" argument of Sampson and Laub.[8] First, it is not simply the occurrence of an event (e.g., getting married or obtaining a job)

[8] Nagin and Paternoster (1993, 1994) make a very similar theoretical argument in their discussion of personal capital and personal control. Nagin and Paternoster argue that the social bond is a developmental "investment" process whereby individuals who make investments in personal capital are, other things being equal, less likely to engage in criminal behavior because of their stake in conformity (i.e., their investment). Nagin and Paternoster note that individuals differ with respect to their inclination to make investments in other persons, institutions, and conventional activities, or in the words of the authors, individuals differ in their "discount rate" (i.e., how they weight present consumption versus future consumption). Yet, the authors argue (and found support for in their study of college respondents) that even individuals with high discount rates can benefit (in terms of reduced criminal activity) by making investments in their social capital (which they call personal capital). This argument is congruent with the argument of Sampson and Laub

per se that serves to reduce the likelihood of crime and deviance, but rather it is the strength of interpersonal ties in the relationship that dictates the negative, neutral, or positive benefits of the relationship (Sampson and Laub 1990, 1997; Laub et al. 1998). Although turning points are frequently envisioned as positive events, negative turning points can redirect a trajectory onto an even more "maladaptive path" (Sampson and Laub 1997; Rutter and Rutter 1994: 244). For example, a male adult criminal who marries a female and has an unstable, conflict-riddled marriage may actually increase the odds or rate of subsequent criminal activity. Second, because social capital is an investment process that develops over time, it is expected that the effects of the investment will also be gradual and accumulate over time. Thus, desistence from criminal activity is better viewed as a developmental process whereby one gradually reduces involvement in criminal and deviant activities over time rather than "going cold turkey" (Sampson and Laub 2001; see also Nagin and Paternoster 1994; Bushway et al. 2001).

Sampson and Laub's theory of informal social control rests on three main themes (Sampson and Laub 1993; Laub et al. 1995). First, informal social controls derived through the social bonds to family and school inhibit delinquent activity during childhood and adolescence, and these two social control mechanisms mediate the effects of background structural (e.g., poverty) and individual factors (e.g., family disruption). During childhood, informal social control largely derives from family processes: monitoring and supervising behavior, consistent application of discipline, and attachment between the parent and child. During the adolescent years, schools are added to list of important social institutions, as well as peer groups and the juvenile justice system.

Second, Sampson and Laub stress the importance of continuity in behavioral tendencies over the life course. Antisocial behavior during childhood and adolescence predicts negative adult outcomes in a variety of life domains (e.g., adult crime, incarceration, frequent unemployment, marital instability). Third, even in the presence of a pattern of stability in behavior across time, salient life events (turning points) associated with social ties to the adult institutions of informal social control (attachment to the labor force, cohesive marriage; military service) can serve to modify or redirect trajectories of criminal offending, *regardless of prior individual differences with respect to criminal propensity*. Stated more pointedly, "childhood pathways to crime and deviance can be significantly modified over the life course by adult social bonds" (Sampson and Laub 1990: 611).

Explaining Continuity & Discontinuity in Crime

According to Sampson and Laub, continuity and discontinuity are the result of two processes. First, they agree with Gottfredson and Hirschi that there are individual differences with respect to criminal propensity and that the self-selection argument cannot simply be dismissed. Thus, Sampson and Laub do agree that part of the observed patterns of continuity result from persistent individual differences in criminal propensity and that low self-control tends to be relatively stable for periods of time (Sampson and Laub 1993: 306, 1997: 155). However, they completely disagree with Gottfredson and Hirschi that persistent individual differences rooted in early childhood are the end of the story. They take exception with Gottfredson and Hirschi's assertion that individual differences in crime propensity persist over time and that social processes and

experiences during adolescence and adulthood have no ability to alter patterns of criminal behavior. Rather, Sampson and Laub argue in support of a process of state dependence, that for better or for worse, is also responsible for patterns of both continuity and change in crime. This is why their theory is best described as a "mixed" theory allowing for both population heterogeneity and state dependence effects (Nagin and Paternoster 2000).[9]

To start, Sampson and Laub (1992: 73, 1993: 21) argue that levels of criminal propensity can change over time as a direct consequence of life events and social processes that modify the strength of one's social bonds. Of course, state dependence is a "double-edged sword" that can serve to either reduce or increase the strength of one's social bonds, and which thus could either increase or decrease the likelihood of future participation in crime and deviance. They posit that criminal involvement at any point in time can weaken or sever the social bond through a process known as *cumulative continuity* (Caspi et al. 1993; Sampson and Laub 1993). Cumulative continuity refers to the process whereby the consequences of behavior at one point in time serve to directly influence both opportunities and behavioral choices at later points in time. "Cumulative continuity is generated by the negative structural consequences of delinquency for life chances" (Sampson and Laub 1993: 124). From Sampson and Laub's (1997) viewpoint, social processes that result from criminal activity, including negative labeling effects, tend to channel individual traits such that people with low self-control often have a diminishing social bond to the social order as a direct *consequence* of criminal activity. Criminal behavior early in life tends to evoke repressive social responses that serve to

[9] To make this point clear, consider the following quote of Laub and Sampson (1993: 306): "the cumulative continuity of disadvantage is thus not only a result of stable individual difference in criminal propensity, but a dynamic process whereby childhood antisocial behavior and adolescent delinquency foster adult crime through the severance of adult social bonds."

further weaken the individual's social bond and make future crime and deviance more likely (Sampson and Laub 1997: 144, 154).

In other words, there is a "closing down" of future conventional opportunities or "closed doors" that leave fewer options for a conventional life (Caspi and Moffitt, 1993; Moffitt 1993). As Laub and Sampson (1993: 306) note, "delinquency incrementally mortgages the future by generating negative consequences for the life chances of stigmatized and institutionalized youths." Over time, the effects of cumulative continuity build into a process known as *cumulative disadvantage* (Dannefer 1987) such that escape from the criminal lifestyle becomes increasingly more difficult.[10]

Yet, Sampson and Laub point out that even though the process of cumulative continuity (state dependence) encourages continuity in behavior, change can and frequently does occur because things can get better (even for persistent chronic offenders) just as they can get worse (Nagin and Paternoster 2000). Thus they contend that desistence from crime can occur to the degree that there is a positive shift in the social bond between a repeat offender and the sources of informal social control, which in adulthood are argued to come primarily from marital cohesion and attachment to the labor market. In other words, qualitative changes in the social bond can occur during adulthood, and the social capital resulting from experiencing positive transitional life events or turning points can help build other conventional relationships that both further strengthen social bonds and simultaneously decrease criminal propensity. Sampson and Laub argue that positive adult experiences can increase an individual's stake in

[10] Cumulative advantage/disadvantage is also referred to as "The Matthew Effect," from the biblical quote, "To him who hath shall be given, from him who hath not shall be taken away that which he hath" (Dannefer 1987)

conformity, as well as provide further opportunities to experience other sources of informal social control. In sum, both stability and change are often embedded in adult life events, which can modify the propensity to engage in criminal and deviant behavior (for better or for worse) despite the level of the individual's prior criminal propensity.

Sampson and Laub have also criticized the casual operationalization of the concepts of "continuity" and "stability." For example, Sampson (1998) questions a recent study of continuity in criminal careers, aptly titled *Continuity and Discontinuity in Criminal Careers* (Tracy and Kempf-Leonard 1996), for defining continuity in crime as anyone with one juvenile arrest before age 17 and at least one arrest between ages 18 - 26. Sampson (1998) finds a contradiction between the authors' conclusion that "continuity was by far the most likely transition" because individuals with an arrest as a juvenile were more likely to be arrested as adults, while they simultaneously report that two-thirds of the individuals with arrests as juveniles were never arrested as adults. Sampson (1998: 1150) notes, "some readers might reasonably interpret this pattern as discontinuity imposed on an aggregate pattern of normative stability" that entirely ignores the amount of within-individual change that actually occurred among two-thirds of the juvenile delinquents. Sampson and Laub (1992) have argued that this concept of "normative" or "relative" stability serves to reify the concept of stability such that there is a misconception about the amount of within-individual change that is taking place over time (see also Cline 1980; Loeber and Stouthamer-Loeber 1998; Sampson 2000). As Sampson (2000: 712) has recently argued, "despite aggregate stability, that is, there is far more heterogeneity in criminal behavior over time within-individuals...change is near ubiquitous."

In essence, the key theme of Sampson and Laub's theory is the "theoretical commitment to the idea of behavioral malleability across the life course and the focus on the constancy of change" (Laub and Sampson 2001: 44-45). This has important implications for the explanation of both the shape of the age-crime curve and the relationship between past and subsequent criminal activity.

Explaining the Age-Crime Curve

Sampson and Laub's theory of age-graded informal social control can be used to explain the observed age-crime curve. Recall that Gottfredson and Hirschi see the general crime decline as occurring for all individuals largely resulting from maturational processes and reductions in criminal opportunities (presumably largely due to aging). Sampson and Laub, on the other hand, see the general decline in crime with age as a result of "institutional forces associated with employment, marriage, prison, and the military that affect bonds to conformity in adulthood" (Cohen and Vila 1996: 144). Thus, the rapidly increasing offending rates in the mid to late teen years can be viewed as a weakening of the social bond as individuals enter adolescence, a period of time when their social bond with their family/parents is strained and they are not yet experiencing the changes in the social bond that generally occur with the positive transitional events of adulthood.[14] As adolescents enter adulthood and experience the informal social control from their investments in marriage, parenthood, and work, crime becomes less attractive due to risks that have accrued through the formation of attachments and commitments of

[14] Tittle (1988) made a similar argument for the shape of the age-crime curve using the social bonding theory of Hirschi (1969) although he did not tie the crime decrease in adulthood to changes in the adult sources of informal social control.

adult life. Thus, the decrease in offending is not due to "inexorable aging of the organism" (Gottfredson and Hirschi 1990: 141), but results from the strengthening of the social bond that often accompanies movement into the various adult roles and responsibilities (Cernkovich and Giordano 2001: 372).

Sampson and Laub (1990, 1992, 1993) are especially critical of the stability and invariance hypotheses of Gottfredson and Hirschi (1990) arguing that such hypotheses are a classic example of the ontogenetic fallacy described by Dannefer (1984). According to Dannefer (1984), the ontogenetic fallacy refers to attributing an outcome solely as a consequence of a preexisting personal trait of the individual rather than recognizing that the outcome is actually the result of interactions between the *social environment* and the personal trait. Dannefer's (1984: 106) argument, directed at biological and psychological models of adult development, was that "sociological research and theory provide the basis for understanding human development as socially organized and socially produced, *not only by what happens in early life*, but also by the effects of social structure, social interaction, and their effects on life chances throughout the life course" (emphasis added). Sampson and Laub (1990: 612) reiterate that many sociological theories of crime (such as pure state dependence arguments) are problematic because they ignore the developmental consequences of the events and processes of early childhood and are excessively fixated on the adolescent years.[12]

[12] Baltes and Nesselroade (1984: 842), in response to Dannefer's (1984) article, criticized sociological theories of adult development for overemphasizing the "intra-individual plasticity (modifiability)" of individuals and for explicitly ignoring the developmental consequences of the "first half of life." Sampson and Laub's mixed theory (allowing for both state dependence and population heterogeneity processes) can be seen as incorporating the critical arguments of both Baltes and Nesselroade (1984) and Dannefer (1984).

Building on the comments of Dannefer (1984), Sampson and Laub (1990: 609) argue that their model is "sociogenic" because it explicitly incorporates not only individual differences, but also acknowledges how salient life events in adulthood play an important role in determining the amount of change in an individual's criminal offending trajectory over time. Sampson and Laub challenge the "invariance" argument because it cuts at the core conceptual foundations of the life course perspective on several fronts, especially the presumption that time and place matter in the lives of individuals. The invariance argument posits that trajectories do not vary even as social conditions change, which is a direct attack on the principle of contextualism discussed above. As Laub and Sampson (2001: 44) note:

> Life-course accounts embrace the notion that lives are often unpredictable and dynamic and that exogenous changes are ever present. Some changes in the life course result from chance or random events, while other changes stem from macrolevel "exogenous shocks largely beyond the pale of individual choice (e.g., war, depression, natural disasters, revolutions, plant closings, industrial restructuring)." Another important aspect of life-course criminology is a focus on situations—time-varying social contexts—that impede or facilitate criminal events.

Explaining the Relationship of Past to Subsequent Criminal Activity

Before moving on to the dual taxonomy theory of Moffitt, a final comment is in order concerning the implication of Sampson and Laub's theory for the relationship of past to subsequent criminal activity. Their use of the concept of cumulative continuity is a state dependence argument, whereby past criminal activity increases the likelihood of subsequent criminal activity as a result of its weakening the individual's social bond to

society. As Sampson and Laub (1997: 144) state, "the state dependence component [of our theory] implies that committing a crime has a genuine behavioral influence on the probability of committing future crimes." Therefore their theory predicts a persistent significant effect between past criminal activity and subsequent criminal activity even after controlling for persistent (unobserved) heterogeneity.

MOFFITT'S DUAL TAXONOMY THEORY

According to the general theories of Sampson and Laub (1993) and Gottfredson and Hirschi (1990), a single theory of crime is applicable to all individuals in the population, and offenders are merely *different in degree*; variations in criminal offending patterns over time are explained purely by variation in the key theoretical constructs of each theory, but the same theoretical explanation applies to all individuals (Dean et al. 1996; Paternoster and Brame 1997; Paternoster et al. 1997). Moffitt (1993, 1997), on the other hand, proposes a *typological* theory of criminal behavior based on the presumption that offenders are *different in kind*, with each "kind" or type requiring a separate, distinct etiological explanation. In the words of Gibbons (1982: 219), the adage "different strokes for different folks" explains the core assumption of any typological theory of criminal behavior. Typological theories of crime have a long-standing history in the field of criminology, but the basis for creating the distinct categories or typologies of offenders has changed from differentiating offenders on the basis of offense type (e.g., property offenders, violent offenders, sex offenders) or skill level (e.g., professional thief versus amateur thief) to more recently differentiating offenders on the basis of broader

behavioral categories and/or longitudinal pathways/trajectories of criminal behavior across the life course (Loeber et al. 1998).

Moffitt's dual taxonomy theory is an example of this recent brand of typological theoretical expositions and was posited by Moffitt as a direct response to the paradox of persistence finding. Moffitt's theory proposes two distinct, unique groups of offenders in the population: life-course-persistent offenders and adolescent-limited offenders.[13] Moffitt argues that each of these offender types follows a distinctly different longitudinal trajectory of criminal/antisocial behavior and that the explanation for each trajectory must use variables that are proximally related to the shape of each offending trajectory. Moffitt proposes one trajectory consisting of individuals who begin offending early in life and then constantly engage in criminal/antisocial activities across adulthood, whereas the other trajectory does not begin offending until the onset of adolescence and then confines or limits their offending largely to the adolescent years (i.e., desists by early adulthood). During the adolescent years, both of these groups are actively offending and Moffitt argues that it is impossible to separate the two groups of offenders using only a cross-section of data; longitudinal data is absolutely necessary to separate out the two distinct groups with qualitatively distinct trajectories of criminal offending (Moffitt 1993, 1997).

[13] Patterson and colleagues (Patterson 1995; Patterson and Yoerger 1993; Patterson et al. 1989, 1993) have presented a theory that similarly divides the population into two groups, labeled "early starter" and "late starter" groups. Similar to the distinction between the theories of Gottfredson and Hirschi (1990) and Wilson and Hernstein (1985), both of these theories present theoretical explanations that have identical longitudinal implications about continuity and discontinuity in patterns of offending, the relationship between age and crime, and the relationship between past and future criminal activity. The differences between Moffitt and Patterson et al. theories surround the precise distal explanation of what causes the existence of each offender type, not that the distinct offender types themselves exist. Moffitt's "life-course-persistent" and "adolescence-limited" offenders resonate with Patterson's "early starter" and "late starter," respectively. Paternoster and colleagues (Paternoster and Brame 1997, Paternoster et al. 1997) have noted that there is a great deal of similarity between the theories even in terms of the theoretical exposition of each offender type but the etiological explanations are not entirely identical.

Since these offender types are argued to be distinct, the causes of their criminal activity must explain why their offending begins and then either persists (life-course-persistent) or desists (adolescent-limited), and the relevant predictors of each type must be proximally related to their predicted offending trajectory (Paternoster and Brame 1997; Paternoster et al. 1997). Like the theory of Sampson and Laub, Moffitt's theory is a mixed theory incorporating both population heterogeneity and state dependence processes; however each process is hypothesized to operate on only one of the distinct offender types (Nagin and Paternoster 2000). In Moffitt's theory, a set of dynamic, state dependence variables is responsible for the offending patterns of the adolescent-limited offenders, whereas a set of static/population heterogeneity variables is responsible for the the criminal behavior of life-course-persistent offenders (Paternoster and Brame 1997).

The Life-Course-Persistent Group

The life-course-persistent group, as defined by Moffitt, accounts for roughly 4-9% of the male population and who (as the name suggests) begin offending early in life (prior to the onset of adolescence/puberty) and persistently engage in criminal/antisocial activities over the duration of the life course.[14] Because of this group's early and persistent criminal behavior, Moffitt grounds her theory of the life-course-persistent

[14] Patterson's theory (Patterson 1995; Patterson and Yoerger 1993; Patterson et al. 1989, 1993) of the "early starter" is more heavily grounded in the effects of poor parenting as being largely responsible for producing the "early starter" group, whereas Moffitt attributes the causes to be neuropsychological deficits combined with poor environmental (family) reactions to it with. Poor or ineffective socialization clearly plays a definitive causal role in both theories. More importantly, both theories agree that by the end of childhood (or the end of the pre-teen years at the latest), this group has been formed and there is little, if anything, that can be done to change their offending propensity thereafter. Criminal offending propensity among this group is thus seen as fixed invariant.

offender type (hereafter referred to as LCP) using factors present very early in life (i.e., proximal to the start of their offending).

The life-course-persistent group consists of individuals who during the early formative years of life are faced with neuropsychological deficits caused by their mother's drug use during pregnancy, poor prenatal nutrition, complications during delivery, pre- and post-natal exposure to toxic agents (e.g., lead), and/or child abuse/neglect shortly after their birth. The neuropsychological deficits leave the "vulnerable and difficult infant" with early deficits in cognitive functioning, emotional reactivity, and verbal and social skills, as well having a generally "difficult" temperament that results in the child being irritable, inattentive, impulsive, aggressive, having poor judgment and low self-control. According to Moffitt (1997: 18), "children with neuropsychological problems evoke a challenge to even the most resourceful, loving and patient families."

Unfortunately, however, these difficult children are generally born into families that do not have the social, psychological, and/or financial resources nor the parenting skills necessary to deal with the unruly, difficult child. Thus, Moffitt argues that these neuropsychological deficits (i.e., personal traits) then reciprocally interact with environmental variables that serve to further exacerbate the child's already difficult personality as a result of being socialized in a criminogenic home environment (Moffitt 1997: 18).[15] Parents with difficult children tend also to have often suffered neuropsychological deficits and difficult temperaments themselves, resulting in

[15] Moffitt's treatment of "criminogenic environments" is limited entirely to the home/family, and neglects to include any insights regarding the contributory role of the macro community setting (e.g., social ecology) to the production of life-course-persistent offenders.

ineffective and counterproductive parenting. Parents of difficult children tend to either overreact and harshly discipline the child for their problem behaviors, to entirely ignore the problem behaviors and the child, or to inconsistently and erratically discipline the child for their problem behaviors.[16] As a result, "the juxtaposition of a vulnerable and difficult infant with an adverse rearing context initiates risk for the life-course-persistent pattern of antisocial behavior. The ensuing process is a transactional one which the challenge of coping with a difficult child evokes a chain of failed parent-child encounters" (Moffit 1993: 682). By the end of the childhood years, the socialization process has broken down and failed, resulting in an individual with a strong, time-stable proclivity to engage in various criminal and antisocial activities (e.g., serious violent offenses, property offenses, drug offenses, sexual promiscuity) across the remaining duration of the life course. The interaction of a problem child with problem parents is a harbinger of the LCP individual.

While the theoretical propositions related to the LCP offender are often characterized as static because of the time-invariant nature of the criminal propensity of such offenders after childhood, it is important to note that the theory is a dynamic theory in the early, formative childhood years (Paternoster et al. 1997). In fact, Moffit draws on the principle of cumulative continuity (i.e., consequences of behavior at one point in time serve to directly influence both illegitimate and legitimate opportunities and behavioral choices at later points in time) to argue that "early individual differences set in motion a down hill snowball of cumulative problems that increase the probability of offending."

[16] It is worth noting the similarity between these parenting conditions and those delineated by both Gottfredson and Hirschi (1990) and Sampson and Laub (1993) as important in their theories.

Moffit argues that the option for future changes in the antisocial propensity of LCP individuals is limited because: (1) they fail to successfully engage in or learn prosocial alternatives to their antisocial behavior as a direct consequence of their neuropsychological deficits that make it extremely unlikely they will perform well in school or prosocial activities (i.e., self-selection), and (2) they become increasingly ensnared in the criminal/deviant lifestyle as a direct result of the consequences of engaging in such activities (cumulative continuity). Interestingly, Moffitt describes the LCP individuals in terms very similar to Gottfredson and Hirschi's description of a low self-control individual, and thus it is not surprising that Moffitt echoes sentiments similar to Gottfredson and Hirschi with respect to the poor odds of LCP individuals changing their behavior over time (Cohen and Vila 1996; Paternoster and Brame 1997; Paternoster et al. 1997). "Simply put, if social and academic skills are not mastered in childhood, it is very difficult to later recover lost opportunities" (Moffitt 1993: 684).

The ineffective socialization of the LCP offenders (a consequence of the interaction of the personal traits and criminogenic environments in which they are raised) produces individuals destined to fail in virtually all aspects of their family, academic, and interpersonal lives who are likely to engage in criminal activities throughout their entire lives. For example, in contrast to Sampson and Laub's proposition that marriage and job ties can decrease the offending propensity of *any* offender, Moffitt argues that LCP individuals will simply select both jobs and wives that serve to support rather than change their antisocial lifestyles (assuming they can relate to jobs and wives at all). In other words, these transitional life events do not function as turning points in the life courses of LCP individuals according to Moffitt. These individuals do not redirect their

criminal/antisocial lives into more conventional ones because they "miss out on opportunities to acquire and practice prosocial alternatives at each stage of development" (Moffitt 1993: 683). Moffitt characterizes the lives of LCP individuals as "maladaptive" because they fail to change in response to any changing life circumstances. Thus, by the beginning of adolescence, the lives of LCP individuals are dominated by a static process that resulted from a dynamic one that began at (or before) birth, and their propensity to engage in criminal and antisocial behavior is hypothesized to be "tenacious across time and in diverse circumstances" (Moffitt 1997: 24). LCP offenders have trouble getting along with individuals in any social setting in which they find themselves, and further, they engage in impulsive, aggressive antisocial behavior in all social settings as children, adolescents, and adults. As Paternoster et al. (1997: 237) accurately describe this group, they "are 'bad apples' who exhibit significant deficits in early childhood socialization and rarely get back on track."

The Adolescent-Limited Group

The second offender type in Moffitt's dual taxonomy theory is the adolescent-limited offender group (hereafter referred to as AL). The AL offender type is the near ubiquitous offender group, and in a statistical sense, their behavior is entirely normal in modern society. Individuals in this offender group only offend for a very short period of time, that is *limited* to the adolescent years.[17] They begin offending in early adolescence

[17] Patterson's theory (Patterson and Yoerger 1993; Patterson et al. 1989, 1993) of the "late starter" is more heavily grounded in the effects of "family disruption" variables, including parental divorce and unemployment. The family disruption variables tend to decrease parental supervision of adolescent activities, which in combination with accessible delinquent peers, provides the social setting for experimental excursions into antisocial and delinquent activity.

(around ages 14-15), commit offenses that are less serious in nature compared to those committed by the LCP group, and have all but ended their criminal offending patterns by the end of adolescence. According to Moffitt (1997: 16), individuals following the AL trajectory of criminal offending "have no notable history of antisocial behavior and little future for such behavior in adulthood."

Importantly, and in stark contrast to the generalized offending pattern of the LCP offender type, this group of offenders engages in "situation-specific" behaviors. In some situations they may well behave in a criminal or antisocial manner, while in other situations they may show no such tendency to engage in such behaviors. According to Moffitt, their tendency to engage in criminal/antisocial behaviors is the result of dynamic variables that represent changes in local life circumstances that shift or alter the reinforcement contingencies (i.e., costs and benefits) of engaging in such behaviors. Given that their offending patterns are hypothesized to be entirely bounded by the adolescent years, the causal factors for this group must be proximal to these ages and account for both the start and stop of their offending patterns. For the AL group, Moffitt emphasizes the importance of dynamic variables that rapidly evolve over a short period of time (the years of adolescence). Moffitt's argument is that changes in these variables lead the AL individuals into starting their offending, and changes in these variables will also be responsible for extinguishing their offending behavior as well.

The AL group of offenders, unlike the LCP group, is hypothesized to lack any underlying, persistent criminal propensity and to have been effectively socialized by their parents. So why do they offend at all? According to Moffitt, individuals in the AL type engage in criminal activity as a result of the strain-inducing *maturity gap* that exists

between biological and social maturity.[18] In all modern societies, adolescents occupy an ambiguous status between childhood and adulthood leaving them in "five-to-ten year role vacuum" (Moffitt 1997: 26). Adolescents, unlike children, are no longer entirely biologically dependent upon their parents; they have reached an age of biologically maturity, and are expected to behave like adults. At the same time, however, they are not given access to adult roles (e.g., work, sex, drink alcohol, marriage) that allow them to enjoy the benefits of behaving like an adult, they are not allowed to make any decisions of real import, and so they cannot experience the social and financial independence of adult life that they increasingly desire. For example, adolescents want to have their own families (or at least engage in sexual activity) and their own places of residence, but they are socially admonished to complete their schooling trajectories prior to beginning their families of procreation or establishing their own housing. As Moffitt (1997: 26) states, "they want desperately to establish intimate bonds with the opposite sex, to accrue material belongings, to make their own decisions, and to be regarded as consequential by adults…[they are] chronological hostages of a time warp between biological and social age."

Eventually, the strain of the cognitive dissonance caused by the maturity gap leaves the adolescents looking for an alternative means to obtain the resource they so desire: mature status. Moffitt (1993: 686) argues that the AL group of offenders engages in a process of *social mimicry* in order to obtain the desired resource:

> Social mimicry occurs when two animal species share a single niche and one of the species has cornered the

[18] Greenberg (1977, 1985) provided a similar "strain" explanation for the age-crime curve.

> market on a resource that is needed to promote fitness. In such circumstances, the 'mimic' species adopts the social behavior of the more successful species in order to obtain access to the valuable resource.

Members of the AL group view the behavior of the LCP offender groups as the embodiment of mature status. LCP offenders resist and fight the powers of authority, they smoke, drink, have sex, are frequently truant from school, often have (stolen) cars, attorneys, and offspring, and generally go about making decisions concerning when and what they will do regardless of what other people want them to do (i.e., act independently). Thus, members of the AL group begin to mimic the behavior of the "more successful species" (the LCP group) in order to obtain access to the valuable resource (mature status). In essence, the AL group emulates the behavior of the LCP group because their lifestyle resembles the experiences of adulthood rather than childhood.

It is important to note that access to delinquent peers is one of the most critical dynamic variables in the explanation of the AL offending pattern, as there must be behaviors to emulate or mimic. Essentially, Moffitt argues that first a few AL boys mimic the behaviors of the LCP individuals, more AL boys join in, and eventually a critical mass is obtained in which offending during adolescence is near ubiquitous. Moffitt argues this access to delinquent peers can be as simple as watching the LCP offenders "at work" from a distance. Mimicry need not involve exchanges of affection (which LCP offenders are presumably incapable of anyway) or actual communication, but rather simply requires the AL offender to (1) "observe antisocial behavior closely

enough and long enough to imitate it successfully" (Moffitt 1997: 29) and (2) be cognizant that the behavior of the LCP resembles adult-like independence.

Thus, Moffitt's explanation of the causes of AL offending is that engaging in criminal/delinquent/antisocial behaviors allows these individuals (who are trapped in the maturity gap) to have access to mature status and all of its resulting power, privilege, and resources. During the adolescent years, delinquent/criminal behavior holds "symbolic value as evidence that teens have the ability to resist adult demands and the capacity to act without adult permission" (Moffitt 1997: 31). The independence and maturity symbolized by delinquent acts is so intrinsically rewarding to the AL individuals that it reinforces the delinquent behavior while they are trapped in the maturity gap. Thus, criminal activity is likely to be followed by subsequent criminal activity in this group because the perceived rewarding consequences of engaging in criminal activity (i.e., obtaining mature status) serves to reinforce the behavior itself.

As members of the AL group near the end of the adolescent years, however, they begin to desist from their criminal offending for two reasons:(1) eventually the perceived rewarding properties of delinquent activities are outweighed by the severe negative costs and consequences (i.e., incarceration) associated with continued engagement in such activities (i.e., the reinforcement contingencies change); and (2) the members of this group begin to obtain access to the valuable resource of mature status through more conventional routes (parenthood, employment) that they did not have access to, but desired, during the adolescent years. Thus, as the AL offenders exit the maturity gap. They begin to desist from committing criminal/delinquent acts because "they realize that

continued participation in crime could threaten their newfound and long-awaited autonomy" (Moffitt 1997: 35).

AL offenders, unlike their LCP counterparts, are able to effectively desist from committing crimes and delinquent acts for three primary reasons. First, unlike the LCP offenders, the AL offenders still have family, occupational, and marital opportunities that they could lose if they continue to engage in criminal activities. Both the informal and formal sanctions associated with criminal activity eventually become too risky for continued engagement in criminal activity.

Second, the AL offenders are assumed to have no underlying criminal propensity, and thus they do not have the usual characteristics associated with that criminal propensity such as poor social skills, poor academic performance, the inability to forge and keep close relationships with other individuals, and low self-control. Because their antisocial and criminal activity did not begin until adolescence, they were able to avoid accumulating these negative personal characteristics and are still eligible for post secondary educational and occupational training opportunities, have good marital prospects, and able to obtain desirable jobs. In short, AL offenders have both good and available options for change, and have the personal characteristics that allow them to take advantage of the prosocial opportunities that become available in early adulthood.

Finally, because their antisocial and criminal activity began at a later age than the LCP offenders, the process of cumulative continuity operated for fewer years for AL offenders. This is especially true since the AL offenders are hypothesized to engage in less serious offenses (e.g., theft, vandalism, drug and alcohol offenses) that do not carry the same severe consequences as the serious offenses that the LCP offenders are

hypothesized to be involved in. Moffit acknowledges that some AL offenders, however, will become ensnared in the damaging consequences of their criminal activities (e.g., incarceration, drug addiction) and that these offenders will have a more protracted period of desistance even into early adulthood.

Explaining Continuity & Discontinuity in Crime

Moffitt began the exposition of her dual taxonomy theory with the explicit goal of accounting for both the shape of the age-crime curve and the paradox of persistence. Moffitt's account for the paradox of persistence (e.g., adult criminal behavior is best predicted by criminal activity during the juvenile years, but most juvenile offenders do not become adult offenders) is that the aggregate mixture of the two offender types is responsible for this finding. One of the offender types, the LCP group, is responsible for the pattern of continuity, while the other offender type, the AL group, is responsible for the change or discontinuity in criminal offending patterns. Since the LCP group begin their offending at an early age and continue offending well into adulthood, "continuity is the hallmark of the small group of life-course-persistent antisocial persons" (Moffitt 1993: 679). The AL group, on the other hand begins offending during early adolescence and desists shortly thereafter, and thus discontinuity is the hallmark pattern of this group. As noted by Moffitt (1993: 691), the differential patterns of continuity and discontinuity between the two groups is the result of the fact that:

> at the cross-roads of young adulthood, adolescence-limited and life-course-persistent delinquents go different ways. This happens because the developmental histories and personal traits of adolescence-limiteds allow them the option of exploring new life pathways. The histories and traits of the

> life-course-persistents have foreclosed their options, entrenching them in the antisocial path.

Since one group is characterized by a pattern of continuity (LCP) and the other is characterized by a pattern of discontinuity (AL), longitudinal data should be able to empirically separate or tease out the two different groups of offenders.

Explaining the Age-Crime Curve

In a manner identical to her explanation of the paradox of persistence, Moffitt argues that it is the mixture of the two hypothesized offender types that makes the age-crime curve assume its observed shape. Indeed, Moffitt argues that her typology "addresses the shape of the curve of crime over age...by drawing attention to two trajectories concealed within the curve of crime over age" (Moffitt 1997: 11-12). The upward surge of the curve results from the increasing participation rates of the AL group, whereas the downward surge results from the patterns of desistance of this group. Given that the AL group is assumed to outnumber the LCP group, their offending patterns are argued to dominate the shape of the curve, while the small number of LCP offenders are responsible for composing the childhood and adulthood offenders in the tails of the curve.

Again, the implicit assumption in this argument is that upon empirically separating the two hypothesized groups, one should find two distinct trajectories: (1) the criminal offending trajectory of the LCP group should be relatively flat because they are hypothesized to engage in criminal activities across the life course at a relatively constant rate (i.e., they do not desist); and (2) the offending trajectory of the AL group should show a strong upward surge at the beginning of the adolescent years, and a similar

downward surge at the end of the adolescent years (and which may extend into young adulthood as a consequence of the AL offenders who become more ensnared in the consequences of their criminal activity). To be clear, Moffitt (1993: 695) is adamant that age is not a predictor of subsequent criminal activity *within* the LCP group (because they engage in criminal activity at a persistent rate), whereas age is a strong predictor of future criminal activity of the AL group (because of their trajectory's bounded dependence on the adolescent years).[19]

Before concluding this section, it is important to reemphasize that one of the key assumptions embodied in the dual taxonomy theory of Moffitt is that the heterogeneity of criminal offending across the life course can be decomposed into two discrete groups of offenders (and two only) with distinctly different age-crime curves. If more than two groups were to be uncovered in a study, this immediately would cast serious doubt on the empirical validity of any theory that advocates a "dual taxonomy" approach to the explanation of criminal offending. In a subsequent section below, we review the results of previous empirical investigations that present evidence on this assumption and these studies all shed considerable doubt on the claim that the aggregate age-crime curve can be adequately and sufficiently decomposed into only two discrete offender groups. This empirical result is very important for two primary reasons. First, Moffitt, in fact, discursively presented her theory by largely drawing on a number of cross-sectional

[19] According to Moffitt (1993: 695), the variables that predict membership in the LCP offender type are "health, gender, temperament, cognitive abilities, school achievement, personality traits, mental disorders (e.g., hyperactivity), family attachment bonds, child-rearing practices, parent and sibling deviance, and socioeconomic status, *but not age*" (emphasis in original). For the AL type, Moffit hypothesizes that "individual differences should play little or no role in the prediction of short term adolescent offending careers. Instead the strongest predictive factors should be knowledge of peer delinquency, attitudes toward adulthood and autonomy, cultural and historical context, *and age*" (emphasis in original).

epidemiological studies and that despite her claim that her theory will account for the shape of the aggregate age-crime curve, she has (to the best of our knowledge) actually presented a longitudinal analysis clearly showing that two groups are adequate to explain criminal offending patterns across the age distribution.

Second, and perhaps more importantly, the dual offender types described by Moffitt are frequently used in empirical applications for both interpreting results and completing analyses on the two "offender groups" after dividing the sample into two groups (which are then labeled LCP and AL) solely on the basis of age of onset alone (see e.g., Dean et al. 1996; Piquero et al. 1999; Scholte 1999; Aguilar et al. 2000; Klevens et al. 2000; Mazerolle et al. 2000; Cernkovich and Giordano 2001; Ge et al. 2001; Piquero and Brezina 2001). If there are more than two offender groups, analyses and interpretations based on this dual taxonomy distinction are not only of questionable theoretical import, but they also are at risk of being potentially misleading. If populations/samples/datasets cannot be neatly and discretely divided into two groups (especially arbitrarily on the basis of age of onset), then completing analyses on two groups (derived on the basis of age of onset) is likely to do nothing other than reify the dual offender categories as if they actually exist in the offender population. In Chapter 7 of this study, we present empirical results from the application of recently developed statistical models that allow one to test this empirical assumption of the dual taxonomy theory (see also Nagin 1999).

Explaining the Relationship of Past to Subsequent Criminal Activity

The dual taxonomy theory of Moffitt has two implications regarding the relationship of past to subsequent criminal activity, one implication for each offender type. First, for the LCP individuals, the correlation between past and subsequent criminal acts should be largely nonexistent *within this group* (Paternoster and Brame 1997). These individuals, as a result of their poor socialization, engage in criminal activities persistently across time due to their time-invariant propensity to engage in such acts. Thus, there is a natural correlation between past and subsequent criminal acts, but it is *spuriously* due to their underlying propensity that predisposes them to consistently engage in these behavioral acts. For the AL individuals, there is an opposite expectation that there will be a strong *causal*, state dependence effect resulting from both the consequences (i.e., ensnarement into the lifestyle) and benefits (i.e., the positive reinforcement contingencies of achieving mature status) of engaging in criminal activity. Thus, the dynamic reinforcement contingencies and possible ensnarement consequences of criminal behavior are argued to increase the probability that such behavioral acts will be repeated again in the future. Paternoster and Brame (1997) point out that Moffitt allows for a possible small state dependence effect for the LCP group due to the potential continuing ensnarement (cumulative continuity) processes during adolescence. They also note, however, that most of the correlation should be almost entirely due to the time-invariant high-level of criminal propensity in this group and that any observed state

dependence effect for LCP group should pale in comparison to the observed effect in the AL group.[20]

THEORETICAL SUMMARY

In this section we provide a brief summary of the main theoretical points of the discussions above. In a nutshell, the theoretical controversy between these three theories boils down to a single question (Cohen and Vila 1996): how stable or inflexible are individual differences in the propensity to engage in criminal/antisocial activities across the life course? Because each theory envisions the stability (or instability) of criminal propensity very differently, each makes different predictions regarding both the relationship between age and crime and the relationship between past and subsequent criminal acts, the questions central to this study.

To Gottfredson and Hirschi, between-individual variation in criminal propensity (i.e., amount of self-control) is sufficient to explain both the relationship between age and crime and the relationship between past and subsequent criminal activity. All offenders decrease their offending over time, and the between-individual differences that exist at any one point in time (around age 8) continue to exist at all other points in time. The shape of the age-crime curve is hypothesized to be robust from person-to-person (i.e., the shape is invariant). According to Gottfredson and Hirschi (1990), the propensity to engage in criminal activities is stable over time; change is only "apparent." The age

[20] Paternoster and Brame (1997) note that Moffitt's theory implicitly denies that there should be a large, significant state dependence effect in the LCP group because this group already has a weak bond to society and because Moffitt provides no description concerning why there should be differential cumulative continuity effects within this group. As Paternoster and Brame (1997: 57) note, "Moffitt provides ample reason to believe that everyone fitting the description of the life-course-persistent offender will experience similar kinds of cumulative continuity."

effect (which applies to all offenders equally) cannot be explained by "impotent" sociological variables like marriage, parenthood, jobs, or education. To be succinct, their viewpoint is that desistence "just happens." Further, controlling for stable criminal propensity (which naturally induces a correlation between offending at any two points in time), the correlation between past and subsequent criminal acts will disappear, as the correlation is spuriously due to population heterogeneity in the distribution of criminal propensity.

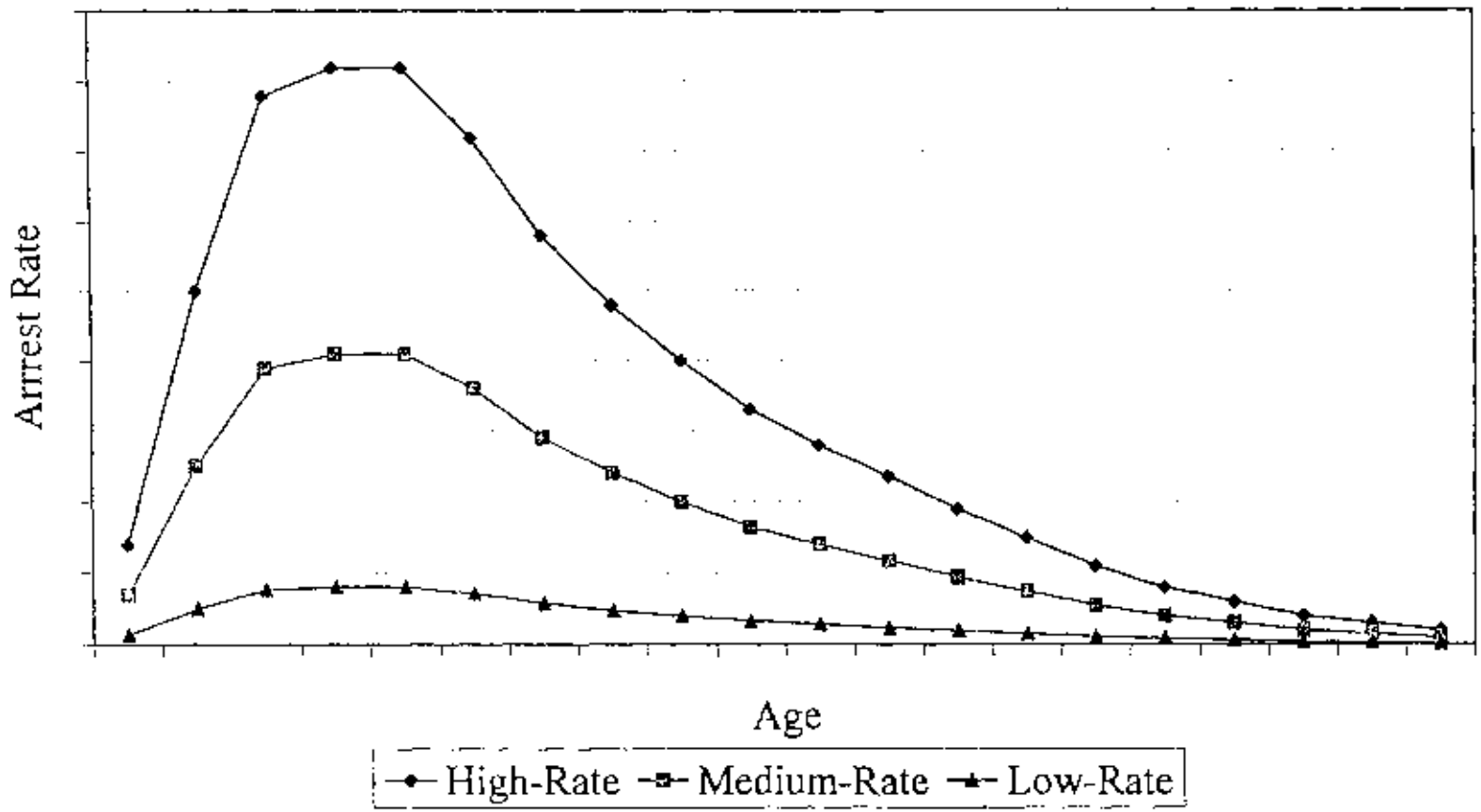
To Sampson and Laub, the relationship between age and crime is due to the varying levels of informal social control across the life course. Adolescence is a period of time when these forces are the weakest (the social bond is weakened during this segment of the life course), but the increasing forces of social control that come with the salient life events of adulthood serve to reduce criminal activity throughout adulthood. It is important to note that Sampson and Laub foresee change as a possibility for all offenders, whether of high or low criminal propensity. The opportunity for change is available for all individuals even though some individuals may not experience change at all, and it may come at later ages compared to others. Sampson and Laub's theory posits that there will be a causal relationship between past and subsequent criminal activity, even after controlling for persistent differences in the propensity to offend, because criminal activity serves to reduce opportunities for prosocial activities and makes continuing in a lifestyle of crime more likely.

Moffitt's dual taxonomy theory envisions patterns of both continuity and change, but each is applicable to only one of the offender types. Change is open to and required from the adolescent-limited offender group, whereas continuity defines the life-course-

persistent offenders. After all, they would not be labeled as "life-course-persistent" if they were expected to desist from criminal activities during their life course. Moffitt also proposes that it is the consequence of mixing the dual offender categories together in the aggregate age-crime curve that is responsible for the observed shape of the aggregate age-crime curve. If one were to separate out the two *hypothetical* groups, one should find two types (and two types only) with radically different offending trajectories. One trajectory should have relatively stable crime rates across time, while the other group's trajectory should have a rapid surge in early adolescence and a similar decline at the end of adolescence. With respect to the relationship between past and subsequent criminal activity, Moffitt implies that the effect should be nonexistent in the LCP group (their offending patterns are largely driven by a failed socialization process during childhood), whereas there should be a strong, *causal* state dependence effect in the AL group (whose offending patterns are largely the result of the "mature status" benefits of criminal activity).

To make the implications of each theory for the relationship between age-crime more concrete, consider the graphical representations of each theory's age-crime explanation as displayed in Figures 2.1-2.3. The "invariance argument" of Gottfredson and Hirschi is presented in Figure 2.1, by three longitudinal offending trajectories—one for high-, medium-, and low-rate offenders—using hypothetical data generated to represent their argument. Each one of the curves follows the inverted-J pattern, and, further, the relative differences between each of the curves is proportional across the age span. The offending rate for medium-rate group is always one-half the offending rate of the high-rate group, whereas the low-rate group's offending rate is one-tenth of the high-

Figure 2.1 Graphical Representation of the Gottfredson and Hirschi Argument Concerning the Age-Crime Relationship, by Offender Type

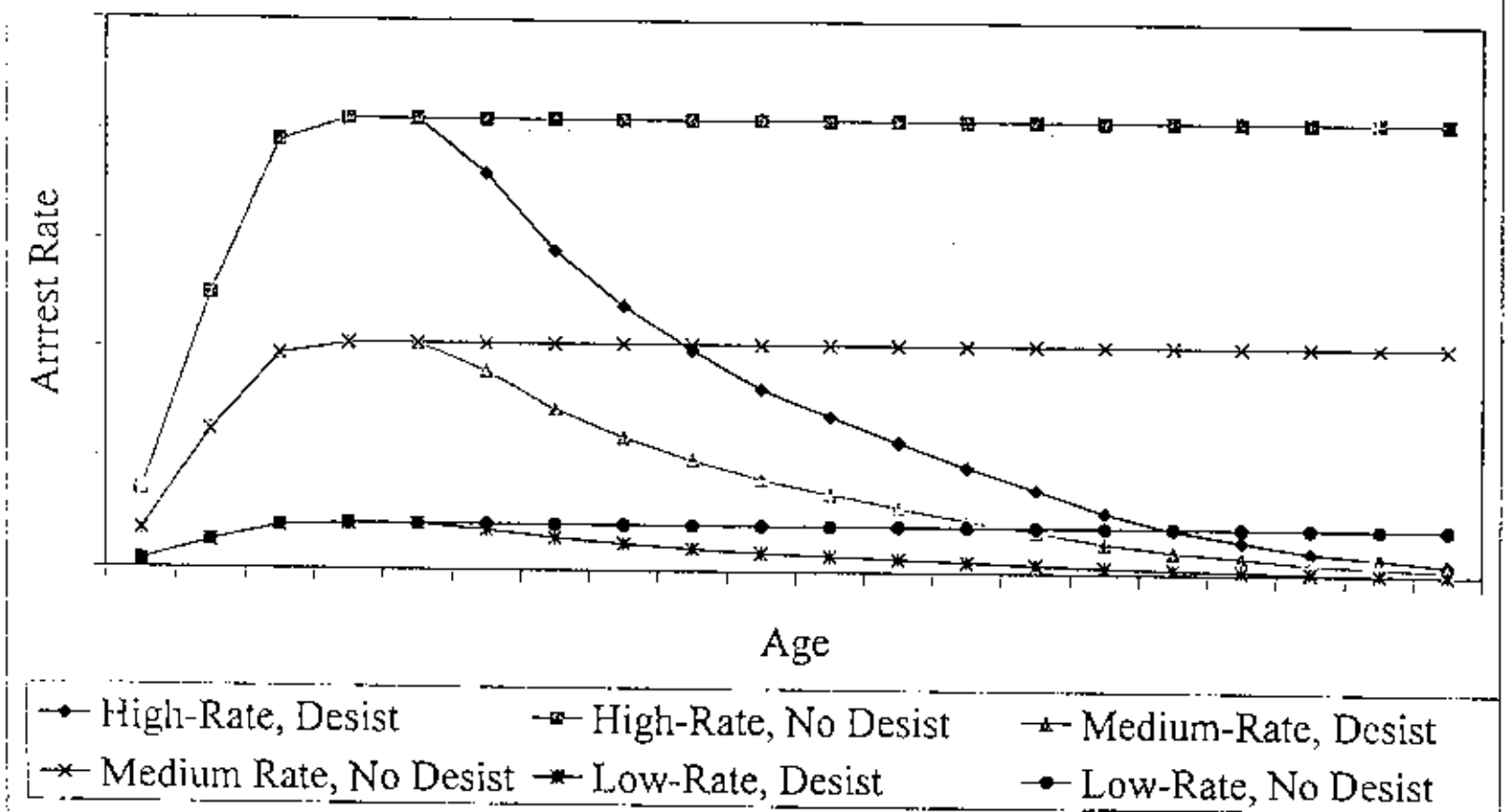


rate group's rate. What causes the differences between the groups is varying levels of self-control, but the actual shape of the curves is identical. The low self-control (high-rate) offenders will start their offending earlier, indefinitely commit offenses at higher rates than the two other groups, and will persist in offending further into adulthood. Thus, varying ages of onset and varying ages at termination from criminal activity merely reflect differences in the distribution of self-control across the population.

Figure 2.2 portrays the argument of Sampson and Laub, only instead of three trajectories as in Figure 2.1, this figure contains six longitudinal trajectories. All three trajectories that appear in Figure 2.1 also appear in Figure 2.2, only now three trajectories that do not display the "decline in crime" with age pattern (i.e., desistence) are also included.[21] For illustrative purposes consider just the two high-rate trajectories. The trajectory that displays a pattern of desistence ("High-Rate, Desist") would correspond to a group of high rate offenders that experienced the salient life events or "turning points" (e.g., cohesive marriages, and stable jobs) during their adulthood. This group of offenders would be theorized to have come under increasing informal social control during adulthood as a consequence of the transitions, and thus their trajectory exhibits a downward pattern during this time (as a consequence of their growing social capital "investment"). The other group, however, would be thought to have missed out on experiencing the key transitional events of adulthood (for a variety of reasons, including

[21] Note that Sampson and Laub do not specifically hypothesize that there are groups with constant rates across time, and in fact they argue that change (especially during adulthood) is near ubiquitous. These trajectories were generated for descriptive purposes only.

Figure 2.2 Graphical Representation of the Sampson and Laub Argument Concerning the Age-Crime Relationship, by Offender Type

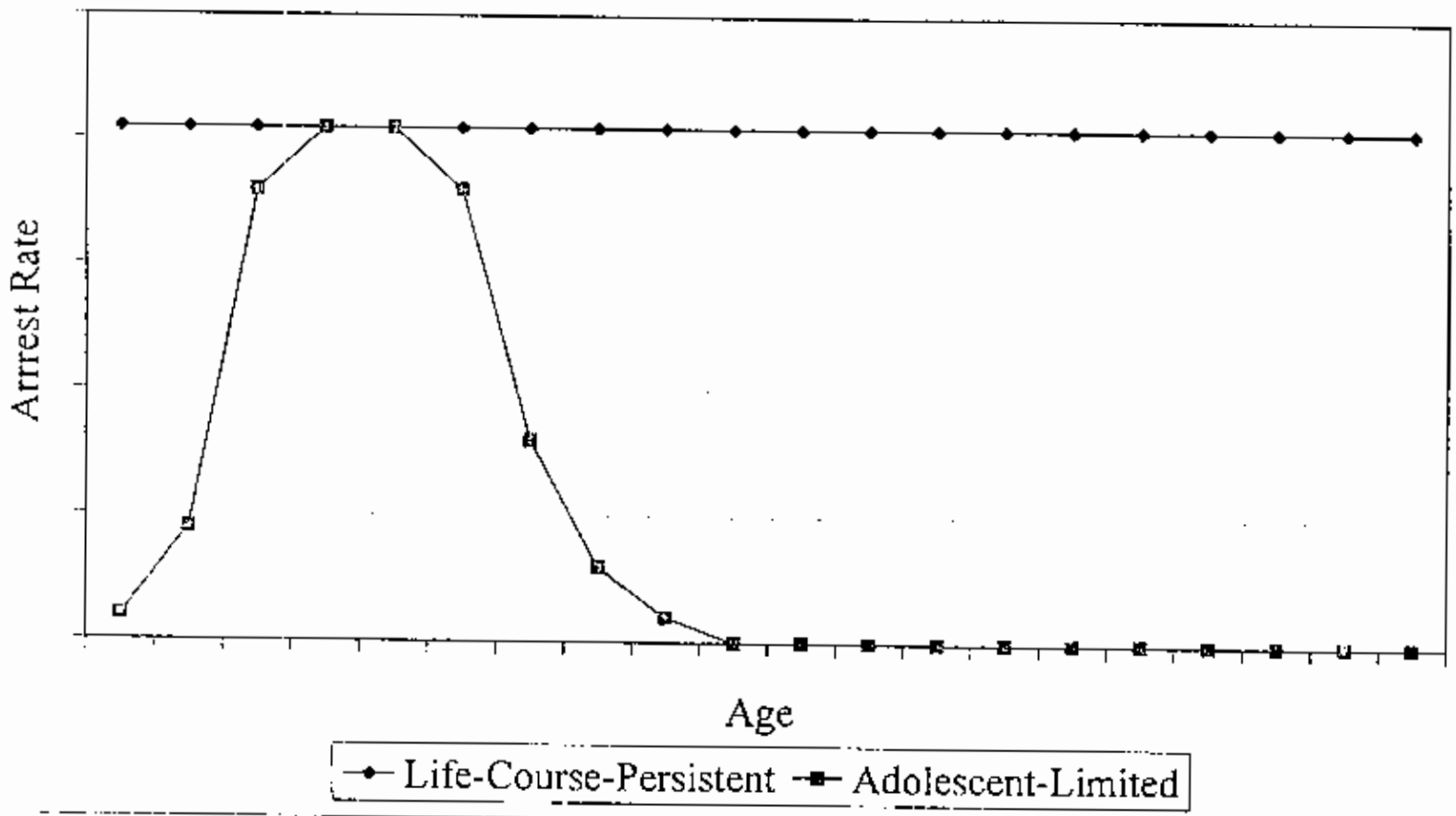


just pure bad luck), and thus they have not experienced the benefits of increasing informal social control during this time; their trajectory simply maintains itself.

Two final points concerning Figure 2.2 are in order. First, the trajectories in this figure were generated to explicitly illustrate Sampson and Laub's notion that change is possible for any offender type, regardless of their prior offending behavior. Even high-rate offenders can experience change. Second, Sampson and Laub posit that adulthood is the precise period of time when preexisting differences become less important than whether or not they experience the salient life events that lead to increasing levels of informal social control. Some offenders will experience the sources of informal social control (i.e., have cohesive marriages and stable jobs), but others will not experience such benefits (in terms of reduced crime) of those sources of social control. In other words, Sampson and Laub would expect significant amounts of change during adulthood that cannot be simply explained away as the mere unfolding of preexisting differences carrying over from the childhood years (Sampson 2000). In Figure 2.2, for example, notice that the "Medium Rate, No Desist" group eventually has a higher arrest rate than does the "High-Rate, Desist" group, even though in the early childhood and adolescent years that group had a significantly higher rate of criminal propensity.

Finally, Figure 2.3 graphically represents Moffitt's argument with respect to the age-crime relationship. The life-course-persistent group maintains a persistent offending rate across time, whereas the adolescent-limited group confines their high-rates offending to the adolescent years. It is important to note that, as argued by Moffitt, the two groups are indistinguishable during the adolescent years, and any cross-sectional data gathered during this time period will not be able to separate out the two groups (nor their unique

Figure 2.3 Graphical Representation of the Moffitt Argument Concerning the Age-Crime Relationship, by Offender Type



etiological explanation of crime). Similar to the implications of the Sampson and Laub explanation, longitudinal data is viewed as absolutely critical for understanding criminal behavior.

PUBLIC POLICY IMPLICATIONS

The previous sections have discussed the theoretical relevance of three particular theoretical perspectives as to the relationship between age and crime (i.e., the stability of differences across time) and the operation of either state dependence and/or population heterogeneity processes. The discussion in this section focuses on the public policy implications of these issues. These issues have profound implications regarding the practical utility of contemporary criminal justice policies for both juvenile and adult criminal offenders. Over the past three decades, the public's reaction to serious crime and the government's response to citizen concerns regarding this social problem have resulted in an increase in both the probability of being sentenced to prison after conviction and the length of time served in custody (Blumstein and Beck 1999; Donziger 1996; Irwin and Austin 1997). This increased punitiveness is the direct result of the "tough on crime" attitude that has swept this country since the late 1960s (Caplow and Simon 1999).

The interaction between the probability of imprisonment and length of time served has led to a phenomenal increase in number of offenders imprisoned over the last thirty years (Blumstein and Beck 1999; Caplow and Simon 1999). Between 1970 and 2000, the adult (prison) incarceration rate in the United States nearly quintupled, increasing from 96 per 100,000 adult residents to 478 per 100,000 in 2000 (Sourcebook

2000). Indeed by midyear 2000, one out of every 142 Americans was incarcerated in either prison or jail, compared to one out of every 218 in 1990 and one out of every 320 in 1985 (Bureau of Justice Statistics, 1995, 2001). In 2000, the total number of adults in the nation's prisons and jails was estimated at 1,931,859 (Bureau of Justice Statistics 2001).

The concern over what has generally been perceived to be a serious crime problem in the U.S., particularly among the young, has become so intense that many states, most notably California, have enacted statutes commonly known as "Three Strikes and You're Out." These laws are proactive crime control policies that mandate the incarceration of individuals who repeatedly commit most of the serious crimes in society—the chronic, or career, offenders. That is, the main stated goal of these programs is to selectively identify those offenders who are often deemed to represent the greatest risk to society, and to remove them from the general public by relegating them to correctional facilities (i.e., selective incapacitation).

In California, the basic content of the "Three Strikes" law requires that defendants with two prior "violent" or "serious" felonies (i.e., those who have already accumulated two strikes), be sentenced to a mandatory term of 25 years to life in prison after conviction of *any* third felony, even if it is nonviolent. Furthermore, this law mandates that any second-strike felony offense receive double the prison time it would receive were it a first offense. Sustained petitions against juveniles, however, do not count as "strikes" under this law until the juvenile reaches the age of 16. Once a juvenile reaches age 16, sustained petitions for the commission of "violent" or "serious" offenses are then counted as "strikes" against them. This fact has the potential effect of increasing the age

at which chronic, youthful offenders are typically sentenced for their third-strike offense. To be sentenced to a period of twenty-five years to life under the "Three Strikes" laws, such chronic offenders would have to: (1) accumulate their first "strike," and serve whatever sentence is accorded them; (2) get arrested and convicted for a second offense, and then serve the time meted out for this second offense; and (3) accumulate a conviction for a third strike offense which then requires a sentence of 25 years to life imprisonment. Thus, it is conceivable that many youthful, chronic offenders will not accumulate their third strike until after they reach the age of 25 because they may be required to serve considerable time periods for their first and/or second offenses after conviction.

There are heavy financial implications associated with incarcerating individuals for extended periods of time. For example, using national statistics on the costs of constructing and maintaining prisons, Irwin and Austin (1997) calculated the cost per additional prisoner (including both supervision costs and the amortized prison construction costs) was estimated in 1997 to be $39,000 per year. In total, the 30-year cost of adding space for just one additional prisoner is estimated to be over $1 million dollars (Irwin and Austin 1997: 139).

Beyond cost, consider the relationship between age and crime and its implications for the use of prisons, especially the draconian policies such as "Three Strikes," as an effective method of controlling crime. If an individual's offending rate is not constant over his or her criminal career, but declines with age, then sentencing older, (previously) high-rate offenders to long prison terms at a point when their offending rates are sharply declining may not be a socially efficient or cost-effective policy (Haapanen

1990; Ezell and Cohen 1997). If all offenders reduce their offending rates as they age, then it is likely that laws such as "Three Strikes" will incarcerate a great number of offenders who, according to some theorists, would appear to present a relatively limited threat to society in the years to come. If, however, there is a group (or groups) of offenders like the life-course-persistent group who commit offenses across the life span at a relatively constant rate, and this group of offenders could be prospectively identified and then segregated from the non-institutionalized population (which is another issue in and of itself), then the rate of serious crime in a society could be reduced substantially (see, for example, Blumstein et al. 1986). The notion that there are offenders who continue to commit crimes at a relatively constant rate independent of their age has considerable seductive appeal from a crime control perspective. It should be noted that much of Gottfredson and Hirschi's initial critique and reaction towards to the criminal career approach to the study of crime was specifically directed at selective incapacitation policies and how these polices completely disregard the relationship between age and crime (Gottfredson and Hirschi 1990; Hirschi and Gottfredson 1986, 1988).

Consider next the implications of the processes of state dependence and population heterogeneity with respect to the practical utility of criminal justice policies including imprisonment to prevent crime. According to the population heterogeneity perspective, criminal propensity once formed *is not malleable*. Thus, from the perspective of Gottfredson and Hirschi (1990), any intervention that is going to have a lasting impact on the criminal propensity of an individual has to involve the efficacy of early child-rearing practices prior to the age of 8 (see also Hirschi 1995). As Gottfredson and Hirschi (1990: 272) pointedly state it,

> Apart from the limited benefits that can be achieved by making specific criminal acts more difficult [e.g., target hardening], policies directed toward enhancement of the ability of familial institutions to socialize children are the only realistic long-term state policies with potential for substantial crime reduction.

Imprisonment, for example, will not have either an enduring beneficial (deterrent) nor negative (criminogenic) effect after release because it has nothing to do with the source of criminal propensity and thus cannot alter it (Nagin and Farrington 1992b). Further, Gottfredson and Hirschi note that any potential beneficial impacts of criminal justice programs are simply a function of selection effects with respect to who got placed in what programs.[22] In other words, high-rate (low self-control) offenders get placed in certain (restrictive, secure) programs, while low-rate offenders get placed in other programs, and success rates of these programs will be entirely dependent on such selection effects.

From the perspective of Sampson and Laub, however, programs that serve to strengthen an individual's bond to society, rather than weakening it, have the possibility to bring about considerable change in the criminal propensity of offenders. Given that the state dependence perspective envisions criminal propensity as malleable across the life course, this perspective envisions the possibility for criminal justice policies to change the probability of future criminal behavior through positive life events (see Laub

[22] This is similar to the problem known in criminology as "net widening" whereby new, less-punitive criminal justice programs often have very low-risk participants placed in them; if it were not for the new program being available, these participants would not have been placed in any type of program specifically because of their low-risk (Walker 1998). Thus, success rates of some programs are sometimes entirely determined by who gets placed in what programs. Studies of "diversion" programs were plagued by this problem (Walker 1998).

et al. 1995). Unfortunately, however, policies regarding positive life events in adulthood are arguably extremely difficult or impractical to implement in practice.

Nonetheless, the current "get tough" on crime policies often simply serve to further isolate the individual and cut off future (positive) opportunities for change. Sampson and Laub argue that imprisonment has criminogenic effects because of its deleterious effects on the prospects of stable employment during adulthood, especially since many of the "escape routes" for convicted felons have been increasingly shut-off as available paths to evade the criminal lifestyle. "Although there is considerable state-by-state variation, licensing boards bar ex-offenders from virtually hundreds of other occupations [besides being a barber], including apprentice electrician, billiards operator, and plumber" (Sampson and Laub 1997: 148; see also Laub et al. 1995). Making prisons even more hostile environments through the removal of educational opportunities, job training, and visitation hours are also not good policies from a state dependence perspective. Reintegrating the offender into society and establishing bonds with conventional persons rather than further isolating them is seen as key to bringing about change in the lives of these individuals (see also Braithwaite 1989). As Laub et al. (1995: 103) write, "it is critical that individuals have the opportunity to reconnect to institutions such as family, school, and work after a period of incarceration."

Having now completed the presentation of the theoretical framework that informs this study, attention in the next chapter is focused on reviewing the previous research on the age-crime and continuity-discontinuity issues.

CHAPTER 3

LITERATUIRE REVIEW AND HYPOTHESES

INTRODUCTION

The theoretical review in the previous chapter sets the stage for a through review of the extant empirical literature concerning the key issues of this study: (1) the relationship between age and crime (and how such a relationship either supports or refutes the stability of differences in criminal propensity over time and the existence of two discrete groups of offenders) and (2) the relationship between past and subsequent criminal activity. The two main sections of this chapter review previous studies that have addressed these two issues. Included at the end of both sections is a discussion of the general findings, the limitations of the prior research, and the hypotheses that guide the analysis of data for this study. This chapter concludes with a discussion of the possible contributions this study can make to the extant literature.

STUDIES OF THE AGE-CRIME CURVE

Given the concerns of Blumstein and colleagues (1986, 1988a, 1988b) and Moffitt (1993, 1997) that age-crime curves aggregated over individuals (i.e., calculated for samples as a whole) may conceal considerable heterogeneity in the offending trajectories of individuals, this review is limited to studies in which the authors have disaggregated their samples into "latent classes" or "latent groups" on the basis of the similarity of their longitudinal offending trajectories. Land (1992) has noted that

Table 3.1 Summary Information of Prior Studies of the Age-Crime Relationship

Authors	Data Source	Sample Size	Gender	Ages Studied	Risk Level[A]	Dependent Variable	Classes Found
Nagin and Land (1993)							
	1961 Cambridge Study	403	Males	10 - 32	Low	11 Two-year Convictions Counts	4
D'Unger et al. (1998)							
	1958 Phildephia Birth Cohort	1000	Males	8 - 26	Low	Police Contact Counts	5
	1942 Racine Birth Cohort	353	Males	8 - 30	Low	Police Contact Counts	5
	1949 Racine Birth Cohort	721	Males	8 - 25	Low	Police Contact Counts	4
	1955 Racine Birth Cohort	1067	Males	8 - 22	Low	Police Contact Counts	5
Laub et al. (1998)							
	1950 Glueck Study	480	Males	7 - 32	High	Arrest Counts	4
McDermott and Nagin (2001)	1977 National Youth Survey	835	Males	11 - 24	Low	8 Years of Self-Reported Serious Offenses	3
Piquero et al. (2001)							
	California Youth Authority Parolees	272	Males	18 - 33	High	Arrest Counts	6

[A] Risk level is defined here as "Low" and "High." Low-risk samples correspond to general population samples that are likely to include a majority of low-risk cases in the data. High-risk samples, on the other hand, is used to refer to samples where high-risk cases will constitute the majority of cases (e.g., samples of parolees).

distinguishing between the various age-crime relationship arguments requires the use of models specified at the individual-level that specifically allow for incorporating controls for heterogeneity in the propensity to offend. The statistical methods available for modeling the presence of separate trajectories have only become available since Nagin and Land (1993) formulated a statistical model consistent with the aforementioned specifications previously noted by Land (1992).

Nagin and Land (1993) introduced the use of semiparametric mixture models to the discipline of criminology as a statistical method able to identify distinct trajectories of criminal offending. Accordingly, all of the studies reviewed below employ the use of the finite mixture methods of Nagin and Land (1993). These finite mixture methods assign each individual to the latent class with the trajectory that most closely resembles the individual's actual observed crime trajectory. Briefly, the mixture methods of Nagin and Land explicitly assume that the sample (population) is composed of a "mixture" of groups, each with their own distinct trajectory, and this modeling strategy both extracts the underlying trajectories present in the data and assigns each individual to the group to which the individual has the highest posterior probability of belonging (Nagin 1999).[1] Table 3.1 summarizes the key information contained in the individual studies reviewed here.[2]

---

[1] In essence, the model fits separate constants and age parameters for each latent class, which allows the shape of each latent class's trajectory to be distinct. See Chapter 5 for further details on the finite mixture methods of Nagin and Land (1993).

[2] Two studies (Fergusson et al. 2000; and Chung et al. 2002) that employ the use of finite mixtures models are not reviewed here. Fergusson et al. (2000) studied the age-crime curve for a sample of adolescents born in Christchurch, New Zealand in 1977. Criminal offending data were only available from ages 12-18, and the authors note that their study thus presents a very limited view of the age-crime curve because both the childhood and adulthood years were truncated from the analysis. Analysis of a binary indicator of police contacts during those ages did uncover four distinct offender groups, including "nonoffenders," "moderate offenders," "adolescent onset offenders," and "chronic offenders." The trajectories of all four groups

As previously mentioned, Nagin and Land (1993) were the first to present evidence concerning both the number of distinct latent classes or offender groups and to discuss the relationship between age and crime for each specific group. In this influential article, Nagin and Land presented their semiparametric finite mixture Poisson model and applied it to the data from the Cambridge Study in Delinquent Development of West and Farrington (1973, 1977), which is a prospective study of 411 males from a working-class section of London that began in 1961 when the boys were 8 years old. The Nagin and Land study used criminal conviction data gathered between the ages of 10 and 32, with 11 time "periods" of conviction counts comprising the dependent variables used in the analyses (e.g., convictions at ages 10 and 11 constituted one "period" of data, 12 and 13 another period, and so on).[3]

Applying the semiparametric mixed Poisson model to these data, Nagin and Land uncovered 4 distinct groups of offenders. The groups were named according to their offending "style": "nonoffenders" (64% of the sample), "adolescent-limited" (12.7%), "low-rate chronics" (9.9%), and "high-rate chronics" (13.4%). The nonoffenders group was comprised of the sample members who had no convictions during the follow-up

---

peaked at age 18, the last year of the follow-up period. The other study by Chung et al (2002) used data from the Seattle Social Development Project, (SSDP) a longitudinal study of male and female youths originally drawn from 18 Seattle public elementary schools. The dependent variable in the study consisted of self-report *offense seriousness* scores (measured at 5 time points between ages 13 and 21). Chung et al. (2002) found a five-class model to provide the best fit to the SSDP data. The five classes included "nonoffenders" (24%), "late onsetter" (14%), "desister" (35%), "escalator" (35.3%), and "chronic" (7%). Since their results do not speak to the issue of trajectories of criminal offending (but rather to trajectories of offense seriousness), however, and given that it is possible for one's offense seriousness score to increase at the same time one's frequency of offending is declining, their results provide ambiguous evidence concerning the relationship between age and crime.

[3] Available data limited Nagin and Land (1993) to using convictions as their dependent variable. The conviction data specifically excluded convictions for traffic offenses and other nonserious offenses (e.g., drunkenness)

period. Obviously, the use of convictions (rather than police contacts or arrests) made it very likely that this category would constitute the largest group in the data set.

Importantly, Nagin and Land uncovered three different groups of offenders within these data, each with their own distinct offending trajectory. They noted that there was considerable heterogeneity in the peak age of offending among the various groups. The peak age of offending for the adolescent-limited group was 14, whereas it was 18 and 22, respectively in the high-rate chronics and low-rate chronics groups. The rate of offending at the peak age (as measured though conviction counts) also varied dramatically among the three groups: with 0.63 convictions for the "adolescent-limited," 1.22 convictions for the "high-rate chronics," and 0.27 for the "low-rate chronics."

Interestingly, their analyses contradicted the assertions of Gottfredson and Hirschi (1990), by finding that between-group age differences in convictions were *not* stable over time. Although the low-rate chronics group did have a peak offending age, their overall trajectory was amazingly flat between the ages of 16 and 30, and the difference in offending rates between the low- and high-rate chronics groups was only 0.15 by age 30, whereas the difference was about 1.0 at age 16. The high-rate chronics group was already highly active in crime at age 10, with this group already having an average conviction rate of roughly 0.8 convictions at that precocious age. This group did, however, show a significant decrease in their conviction patterns (after their peak rate at age 18) as they progressed through adulthood, a finding that is consistent with the assumptions of both Sampson and Laub (1993) and Gottfredson and Hirschi (1990). By finding a low-rate chronics group, Nagin and Land were the first researchers to offer

empirical evidence of considerably more heterogeneity than the two subgroups posited by Moffitt's dual taxonomy theory (1993, 1997).

D'Unger and colleagues (D'Unger et al. 1998) conducted the most extensive examination of the age-crime curve to date when they analyzed five separate datasets.[4] One set of data pertained to the same set employed in the Nagin and Land (1993) study, and since the results obtained in these two studies are identical here, they are not discussed. The four new sets of results presented by D'Unger and her colleagues include analyses of data from the 1958 Philadelphia Birth Cohort study (Tracy et al. 1990), and the 1942, 1949, and 1955 Racine birth cohort studies (Shannon 1988, 1991).

The 1958 Philadelphia Birth Cohort study longitudinally tracked all 13,160 males and 14,000 females born in Philadelphia in 1958 *and* who resided in the city through their 18th birthday. The frequency of "police contact" for felony and/or misdemeanor criminal offenses was collected through age 26 from Philadelphia Police Department records. Police contacts include both actual arrests by law enforcement personnel as well as law enforcement "contacts" that were handled "remedially" or "informally" (e.g., released at the scene or released to parents) and did not involve a formal arrest where the individual is taken into custody (Tracy and Kempf-Leonard 1996). For computational reasons, D'Unger et al. (1998) estimated their models on a random sample of 1000 males.[5]

---

[4] The main purpose of the D'Unger et al. (1998) study was not to examine the age-crime curve per se, but rather it was to determine when the optimal number of latent classes has been extracted from a dataset. The results of their analyses, however, provide evidence that bears directly on the age-crime relationship.

[5] Estimation times of finite mixture models tend to increase greatly with sample size (Vermunt and Magidson 2000). Also, to keep the results comparable to those obtained by Nagin and Land (1993), only males were included in these analyses. D'Unger and colleagues (1998) compared their results from the 1,000 member sample to those with samples of 500 and 2,000 and report the results to be similar across sample sizes

D'Unger et al. (1998) found a five class or group model to be the best fit to these data, and named their classes by the nature of the respective offending trajectories of the groups. The largest group was labeled "nonoffenders" (comprising 61% of the sample), and the other groups included a "high-rate adolescence peaked" group (1%), a "low-rate adolescence peaked" group (9%), a "low-rate chronics" group (21%), and a "high-rate chronics" group (8%).[6] Interestingly, the four groups who engaged in some level of offending bifurcated into high- and low-rate versions of adolescence peaked and chronic types that tracked each other fairly well over time. Although the rates of increase and decrease in offending with age varied both within and among the "adolescence peaked" and the "chronics" offender groups, each group showed a decrease in offending throughout adulthood. The peak ages of offending were 16 for the adolescence peaked groups, and 18 for the chronic groups. The offense rates peaked at 1.0 (for the low-rate adolescence peaked group), 3.3 (for the high-rate adolescence peaked group), 0.21 (for the low rate chronics group) and 0.95 (for the high rate chronics group). By age 26, however, only the high-rate chronics group still had a non-zero offending rate, and their rate at that age was roughly one-quarter of its peak rate at age 18.

The Racine Birth Cohorts longitudinally tracked the offense histories of all individuals born in Racine, Wisconsin in 1942, 1949, and 1955. For research with these datasets, the dependent variables were the number of police contacts for felony and misdemeanor criminal offenses between the ages 8 and 30 (1942 cohort), ages 8 and 25 (1949 cohort), and ages 8 and 22 (1955 cohort). To make these findings comparable with

[6] D'Unger et al. (1998) refer to the group as "adolescence peaked" rather than "adolescent-limited" because their offending patterns included ages outside the adolescent years of 13-17. Both of these trajectories did, however, display a strong peak during adolescence.

the prior research discussed above, the authors limited their analyses to the white and black male members of the samples. This resulted in final sample sizes of 353 (for 1942), 721 (for 1949), and 1067 (for 1955), individuals respectively.

For the 1942 birth cohort, D'Unger et al. found the best-fitting model to have five distinct offender groups. These groups included a "nonoffenders" group (34.6%), an "adolescence peaked" (20%) group, a "low-rate chronics" (31.4%) group, a "high-rate chronics" group (8.8%), and a "late-onset chronics" group (5.1%). Unlike the previous findings, one offender group was located in this dataset that actually increased their offending with age (the late onset chronics group), with the peak rate of offending for this group occurring at age 28 where it then stabilized through the end of the follow-up period. At age 16, this group had an offending pattern that virtually tracked the adolescence peaked group. At that point, however, the two trajectories diverged with the late-onset chronics group continuing to escalate their offending behavior, while the adolescence peaked group began to desist from offending. Interestingly, the high- and low-rate chronics groups differed substantially in their offending rates between ages 16 and 22 (with their offending rates differing by about 1.0 police contacts per year). By the end of the follow-up period, however, the offending rates of these two groups were nearly identical. The high-rate chronics group experienced a significant decline in offending in early adulthood, whereas the low-rate chronics group was observed to have exhibited a much slower rate of decrease in their offending rate.

A four-class (or group) model for the 1949 sample provided the best fit to this dataset. The group trajectories found by D'Unger et al. for this sample included a "nonoffenders group" (35%), "high-rate chronics" group (peak age = 18; peak rate = 2.1;

5% of sample), a "low-rate adolescence peaked" group (peak age = 18, peak rate = 0.25; 40% of sample), and a "high-rate adolescence peaked" group (peak age 18; peak rate = 0.75; 19% of sample). By age 25, both the low-rate and high-rate adolescence peaked groups had virtually desisted from offending (as measured by police contacts at least), whereas the high-rate chronics group was still experiencing roughly 1.5 police contacts per year at this period in their lives. It is interesting to note that this is the only dataset for which the trajectories generally followed the proportional changes across time argument proposed by Gottfredson and Hirschi (1990).

For the 1955 cohort, a five-group model was found to provide the most accurate fit to the data. The five groups included a "nonoffenders" group (44.5%), an "early-onset adolescence peaked" group (2.2%), a "late-onset adolescence peaked" group (15.4%), "low-rate chronics" group (30.1%), and "high-rate chronics" group (7.8%). Unlike the results from the Philadelphia Birth Cohort data, however, the adolescence peaked trajectories did not neatly bifurcate into simply high- and low-rate versions; they differed greatly on their age of onset as well as their rate of offending. Also, the crime trajectories of the various groups did not remain proportional; rather the rate of change of the trajectories varied considerably between groups. For example, at age 8 the early onset adolescence peaked group and the high rate chronics were vary similar in offending rates. At age 16 their trajectories differed by about 2 arrests per year, and then by age 22 they were nearly identical again. Similarly, the trajectories of the low-rate chronics and the late-onset adolescence peaked groups were identical until age 15, at which point the adolescence peaked group had a surge in offending while the offending by the low-rate chronics held fairly steady thereafter. By age 22, the late-onset adolescence peaked

group had decreased their offending back to a level near that of the low-rate chronics group.

The results of the studies reviewed so far have shed doubt on the assertion of Gottfredson and Hirschi that there is a relative stability of between-group differences in offending across time as well as the contention by Moffitt that there are only two discrete groups of offenders in the population. There appears to be considerable heterogeneity among the various distinct offender groups with respect to both the peak ages and rates of offending, as well as the amount and nature of the changes in the offending rates across the age distribution. The data analyses have also consistently uncovered more than two discrete groups of offenders among those sampled. Next we turn our attention to the first analysis of the longitudinal offending patterns among discrete offender groups within a "high-risk" sample.

Laub et al. (1998) conducted an analysis of the longitudinal offending patterns of the 480 delinquent boys from the original Glueck study of the criminal careers of delinquent boys in Boston (Glueck and Glueck 1950, 1968). All 480 boys were white and all had appeared in the Boston Juvenile Court in the late 1930s. The Gluecks followed the boys into adulthood until the age of 32. Sampson and Laub (1990, 1993; Laub and Sampson 1988) subsequently reconstructed these data and put them into machine-readable form, and then subsequently used it in developing and testing their theory of informal social control. In this study, Laub et al. (1998) used the finite mixture

methods of Nagin and Land to ascertain if there were distinct groups of offenders even within this select group of chronically delinquent boys.[7]

For this study, the dependent variable was the count number of arrests at each age between the ages of 7 and 32. Laub et al. (1998) found that allowing for four distinct groups (or trajectories) provided the best fit to these data. Since all of the members of this sample were official delinquents, there was obviously not a "nonoffenders" group in this dataset. There was, however, a significant amount of heterogeneity even in this select sample of juvenile delinquents. Further, even though all four of the offender trajectories were very similar in their offending rates up through age 13, from that point on there was significant variability in the shape of each group's crime trajectory.

"Group 1" consisted of a high-rate chronic group that had an observed peak offending age of 18 (at about 3 arrests per year), and thereafter their trajectory was relatively constant until the late 20s when their offending rates began to decline. Only 11 individuals in the sample were assigned to this group. "Group 2" was a more moderate chronic offender group, with a peak offending rate of about 1.2 arrests per year at around age 18. This group comprised about 19% of the sample. The offending rate of Group 2 was relatively constant during their 20s, and began to decline by the end of the follow-up period (ages 30-32). Group 3 exhibited an offending pattern very similar to Group 2 through age 16, but then experienced a significant decline in the offending rate over the remaining age distribution curve. By age 32, this group had a negligible offending rate,

[7] In subsequent analyses in that article, Laub, Nagin and Sampson (1998) also used the resulting latent class indicators as a method of controlling for persistent individual differences in models testing the crime preventive benefits of a cohesive marriage. The results of their analyses indicated that, even after controlling for unobserved heterogeneity in criminal propensity, a cohesive marriage was a critical factor in the desistence process. Consistent with their theory, the benefits of a cohesive marriage accrue gradually over time as the investment process accumulates social capital.

whereas Group 2 was still offending at about .8 arrests per year at this age. Group 4 was the group with the lowest offending rate (peak offending rate was .7 at age 16). This group (Group 4), which comprised about 31% of the sample, also had a rather negligible offending rate (of about 0.1) by age 26, where it continued to hover for the remaining 6 years of the follow-up period. The results of this study should be viewed with caution, however, because the subjects in the Glueck data set were not randomly selected from the population nor randomly selected from juvenile court cases (see Cohen and Vila 1996). Because the results of this study are based on a matching sample that was drawn by convenience from the Boston juvenile court records, the generalizability of the results are uncertain.

Using a national probability sample to avoid possible sampling bias, McDermott and Nagin (2001) studied the self-reported offending patterns of the 835 male respondents in the National Youth Survey. The segment of the age distribution studied ranged from 11 through 24, but fewer than half of the respondents were available for sampling at ages 11-13 and 20-24.[8] Therefore, the lack of available data for estimating these segments of the age-crime curve demands caution when interpreting the reported results. The dependent variable was a count of self-reported involvement in rape, auto theft, theft of goods worth more than $50, purchasing stolen property worth more than $50, and breaking an entering.

McDermott and Nagin (2001) found a three-class model to provide the best fit to these data. "Group 1" was engaged in offending from ages 11 through 20 at a relatively

[8] Their analyses covered a period of 8 "age years," with the actual ages studied varying between the respondents depending on the age of the respondent at the first wave of interviews. Although the analyses spanned an 8 year period, there were actually only 6 measurement periods used in the analyses due to unequal spacing of the last interview (which occurred three years after the fifth wave survey).

constant rate (between 1.0 and 1.3 offenses), at which point their offending patterns were found to decline. The offending rate of Group 1 peaked at age 15. It should be noted that 748 individuals (or 89% of the sample) in their analyses was assigned to Group 1, and that most of these individuals reported no criminal activity at each age measured. "Group 2," comprising about 6% of the sample, had a peak age of offending at age 11 with 20 offenses per year. Thereafter, this group showed a significant decline in the offending rate through age 19, at which point the rate leveled off at around 5 offenses per year through age 24. The offending pattern by "Group 3" was nearly antithetical to the pattern observed for Group 2. Group 3, which contained 5% of the sample, showed a precipitous increase in their offending from ages 11 through 21, where the offense rate peaked at a rate of 30 offenses per year. Their rate declined to 23 offenses per year by age 24.

In still another study, Piquero et al. (2001) present an analysis of the age-crime curve for a sample of high-risk cases. This study involved an analysis of the adult offending patterns of a sample of 272 parolees released from the California Youth Authority between 1960 and 1970. This is the same youthful offender correction system from which we analyze data in this study, although the data gathered by Piquero et al. predate the large increase in violent offending in the state of California in the 1980s and did not constitute a random sample of CYA wards. The 272 parolees in this study were "older, had more serious commitment offenses, and/or were uncooperative in other CYA institutions" (Piquero et al. 2001: 57). These youthful offenders were paroled at age 18 from the CYA, and were then followed for 16 years until age 33. The dependent variable was a count of arrest events between the ages of 18 and 33. Thus, while their study concerns a limited segment of the age-crime curve (i.e., adulthood only), it is important

for inclusion in our review because they found considerable heterogeneity in the adult offending trajectories of this select sample of offenders.[9]

In fact, Piquero et al. (2001) found that a model allowing for 6 distinct trajectories provided the best fit to the data. Importantly, this study found a 6-class model to fit the data with and without controls for "exposure time" (i.e., time not incarcerated), although the authors noted that the scale of the arrests trajectories, especially during their early 20s was affected by controlling for exposure time. In the nomenclature of criminology, exposure time is referred to as "time on the street" and is used to control for the amount of time spent incarcerated. When a person is incarcerated, they are incapacitated and cannot victimize the general non-institutionalized population simply as a consequence of their isolation and not because of a change in their motivation to commit criminal behavior (i.e., they are denied the opportunity to victimize the non-institutionalized citizenry).[10]

For the models without a control for exposure time, "Latent Class 1" (18% of the sample) was a group that increased their offending through age 21, which was their peak age of offending (at 2 offenses per year). This group then decreased their rate of offending through age 33; their offending rate was negligible from ages 29 onward. "Latent Class 2" (21% of sample) displayed a trajectory very similar to Latent Class 1,

---

[9] As we detail in Chapter 4, from the mid 1970s onward, the California Youth Authority has consistently been responsible for housing the most seriously delinquent 5% of the youthful offender population in California. Thus, any sample of individuals released from the California Youth Authority is by definition a highly select sample. However, it is also a segment of the offender population that is often of greatest concern due to their serious nature and persistence of their offending. See Chapter 4 for more details on the California Youth Authority and the types of youth who come under its custody.

[10] For example, someone who spends 12 months "on the street" and is arrested one time is very different from an individual who spends 1 month on the street and is arrested one time. There is a difference of 11 months of "exposure time" between these two individuals. This is the logical reasoning behind accounting for "time on the street."

only their offending rate was roughly twice the rate of Latent Class 1. Their rate peaked at age 21 and the decreased thereafter, but they still had a positive offending rate at age 33 (at about .75 arrests per year). "Latent Class 3" (7% of sample) peaked their offending during the first year after release (age 18), had a small decrease in their offending rate during the early 20s, and then had a relatively constant rate (at about 3 offenses per year) through the remaining ages in the distribution. "Latent Class 4" displayed a trajectory very similar to Latent Class 3, albeit at a lower rate than Latent Class 3, from ages 18 though 28. This groups' (18% of sample) arrest rate also peaked at age 18, then held at a relatively constant rate through age 28 at about 1 offense per year, and then displayed a decreasing offense rate through age 33. The fifth latent class (24% of the sample) also had a peak rate of arrest at age 18 (at 2 offenses per year), then displayed a decreasing arrest rate through age 25, at which point this rate became stable at about .5 arrests per year. Interestingly, the sixth latent class (10%) had a very small offense rate at age 18 (about 0.25 arrests per year), but then essentially desisted entirely over the remaining ages. In other words, this group was able to essentially remain arrest free after parole. It is interesting to ponder if this group consists of an adolescent limited group who displayed a high offending rate during their adolescent years but then was able to remain arrest free during adulthood.

Two important differences arose after allowing for an "offset" or control for exposure time in the Piquero et al. study. First, the predicted arrest count was of greater magnitude for both Latent Class 2 and Latent Class 3. Latent Class 2 peaked 2 years earlier than Class 3 at a rate of 7 arrests per year, but then exhibited a pronounced declining arrest rate through age 25, at which point the trajectory assumed the same shape

it did in models without controls for "time on the street." Latent Class 3 also did not experience its decline in the early 20s, but rather this group's arrest trajectory held rather constant over most of the remaining age distribution (at a rate of about 7 arrests). During their early 30s, the group had a decrease in offending of about 1 arrest per year. For the most part, however, the overall shape of those two curves did not vary dramatically between the two models. A second difference among the groups noted by Piquero et al. (2001) was that Latent Class 4 also did not experience a decline in offending in the early 20s or a further decline in the 30s, but rather this group had a constant arrest rate over the entire age distribution (at about 3 arrests per year). The remaining three classes had trajectories that were essentially identical in both models. Piquero et al. (2001: 68) concluded that, "the general shape of the arrest trend appears to be robust to controls for exposure time." The percentage of cases assigned to each latent class was virtually identical across the two models as well. Piquero and his colleagues report that more than anything else, it was the magnitude of the arrests scale that was affected the most by controlling for street or exposure time. Nonetheless, the results from both of their models indicate that there is significant heterogeneity in the adult offending patterns of these serious offenders, but how that adulthood heterogeneity related to prior existing differences could not be determined with these data. Indeed, the results of this study also leave one wondering if the findings would have changed had they access to either the juvenile arrest histories of the sample or to a much larger, random sample of youthful offenders. With an overall sample size of only 272 cases that were not randomly drawn the generalizability of the findings from this study must be viewed cautiously.

We have now completed our review of the prior studies that have addressed the age-crime relationship *within* discrete offender groups (that are internally homogenous with respect to their offending patterns across time). In the next section we place the results of these studies into perspective with a discussion that focuses on both the significant themes that have emerged and the methodological limitations of these prior studies.

Discussion and Hypotheses Related to the Age-Crime Relationship

A general summary of many of the studies we have reviewed here can be found in the first, and arguably definitive, study concerning the relationship between age and crime within distinct offender groups by Nagin and Land. In this study, Nagin and Land (1993: 358) noted, "our findings point to large variation across the population not only in offending levels by age but also in the trajectory of offending over age." The results of their study are illustrative of several themes in the literature particularly relevant to our study.

First, there appears to be a considerable amount of individual variation in the offending rates of individuals. This heterogeneity in offending propensity has been documented across a variety of different settings, including birth cohorts from a small Midwestern town such as Racine to a large urban city such as Philadelphia (D'Unger et al. 1998), in a random sample of the general population (McDermott and Nagin 2001) to samples of the serious youthful offender population (Laub et al. 1998; Piquero et al. 2001), in a sample that uses self-report data (McDermott and Nagin 2001) to samples that use official data (Nagin and Land 1993; Piquero et al. 2001), and across varying cultural

settings such as England (Nagin and Land 1993) and New Zealand (Fergusson et al. 2000). The generalizability of the finding of heterogeneity in individual offending is extremely important because samples are often treated as if one trajectory or group is present in the data and as if the effects of persistent heterogeneity are trivial. Such short-sightedness by researchers can lead to misleading and erroneous conclusions (D'Unger et al. 1998; Land et al. 2001; Maltz 1994; Moffitt 1993, 1997).

A second theme in this literature concerns the significant amount of between-group heterogeneity displayed with respect to the direction and nature of change in the shape of the crime trajectories across age. Given that the theories of Gottfredson and Hirschi (1990) and Moffitt (1993, 1997) predict neat, clearly defined changes in offending trajectories over time, while Sampson and Laub (1993) predict more heterogeneity in crime trajectories over time (especially in the adult years), the evidence at this point appears to favor the theoretical position of Sampson and Laub. For example, the McDermott and Nagin (2001) study found a crime trajectory that continued to increase across age, while D'Unger et al. (1998) discovered high- and low-rate "chronics" display markedly slower change in their crime trajectories in comparison with the adolescence peaked groups in their data. Laub et al. (1998), report crime trajectories in their sample that were quite similar at early ages to show markedly differential growth patterns during adulthood. Because the theory of Sampson and Laub predicts greater heterogeneity in crime trajectories than does the other two theories, it appears to receive more support from the previous research.

A third theme in this literature involves two trajectory "regularities" in many of the studies reviewed here. The first regularity is that two distinct primary groups have

been uncovered across many of the studies: (1) the "chronic offender" group where crime peaks between ages 17-21 and then drops slowly in 20s and (2) the "adolescent peaked" group where crime peaks between ages 15-18 and then drops rapidly to near zero by age 22 (D'Unger et al. 1998). The second consistent pattern is a common crime trajectory shape that often bifurcates into high- and low-rate groups that track each over the age distribution (D'Unger et al. 1998).

Yet, even in the face of these regularities, it should be noted that neither the longitudinal shapes nor the number of distinct trajectories were entirely consistent across the various studies. For example, several crime trajectory patterns have only been identified in one or two of the studies, most notably the "late onset chronic" offender found in D'Unger et al.'s (1998) analysis of the 1942 Racine Birth Cohort. While all of the studies uncovered more than two discrete groups, the exact number of classes has ranged from three in McDermott and Nagin (2001) study to six in the Piquero et al. (2001) study. Most studies report identifying four or five distinct crime trajectories. Notably this finding directly contradicts Moffitt's hypothesis of two distinct offender trajectories and seriously questions of whether two trajectories are sufficient to capture the variation of offending trajectories in the population. If there are not just two distinct offender trajectories, then how many are sufficient? Does the number of crime trajectories identified depend on the sample composition? How stable are the identified latent classes within a given population over time? While definitive answers to these questions remain for future research, results such as those presented in Chapter 7 of this study can expand our understanding of these issues by examining the three different samples of serious youthful offenders to be used in our study.

The current literature of the age-crime curve for distinct groups of offenders has several limitations that highlight the need for further study. First, some of the previous studies have focused on rather limited segments of the age distribution (due to limitations of the data sets), with several studies not beginning their measurement of offending behaviors until the onset of adolescence or later, and most of the studies ending their follow-up periods prior to or around the age of 25. The study of Laub et al. (1998) has the longest follow-up period to date, examining the nature of the offending trajectories of 480 delinquents from age 7 through 32. The nature and shape of offending patterns beyond the early thirties are currently not well understood. Second, the study of D'Unger et al. (1998) is the only study to compare the results from datasets generalizable to the same population over time. This makes it very difficult to replicate not only the existence of a crime trajectory group over time, but also whether there are any changes in the precise number or nature of the offending trajectories over time. As such, D'Unger et al. (1998) argue that replication of offending trajectories is a critical research need that is necessary to prevent reifying any particular identified offending trajectory as a stable element in a population. As D'Unger at al. note (1998: 1624-1625),

> The effects of age, cohort or sample composition, and historical setting all play important roles in influencing individual development, hence the variation in trajectories over time. Social context must be viewed as a 'force in development' (Elder and O'Rand 1995), which has the power to alter trajectories of myriad types of behavior.

A final limitation of the previous research mentioned here concerns the analyses of the "high-risk" samples; only two studies have focused on select samples of "high-risk" offenders. Both of those studies, however, have limitations that require additional

research on this critical segment of the offending population. The Laub et al. (1998) study was based on the offending patterns of a sample of white, male delinquents from Boston measured from the 1930s to 1960s, and thus a key question is whether trajectories similar to the ones they describe can be found in more contemporary samples of the population. This is especially significant given that the nature of criminal offending appears to have changed dramatically (i.e., became more violent) after the point in time when their data were gathered. Piquero et al. (2001), on the other hand, only had access to data regarding offending patterns of a sample of serious youthful offenders during their adult years (until age 33). Data from these subject's juvenile years were entirely absent from their analyses. This limitation raises several interesting questions with respect to this segment of the population: (1) how do differences in offending trajectories during the juvenile years relate to the nature of offending during the adult years; and (2) is there an adolescent-peaked group within this population?[11] Furthermore, both of these studies were based on comparably "small samples," and thus we wonder to what degree their findings (or a particular latent class) are a consequence of sampling variation? This question becomes more interesting once we consider that neither of the samples *were* randomly drawn. Thus, it is our contention that there is a critical need for subsequent empirical investigation of the nature of offending trajectories within the population of serious youthful offenders, a contention that has been echoed by Laub and Sampson (2001), Scholte (1999), and Tolan and Gorman-Smith (1998).

[11] Recall that Piquero et al. (2001) found a group with an offending trajectory that by age 20 had terminated their criminal activity (as measured by arrests).

Given the findings and limitations of the literature discussed above, this study will investigate four hypotheses related to the age-crime curve using three relatively large, random samples of serious youthful offenders (to be described in greater detail in Chapter 5):

H_1: There are multiple groups or latent classes of offenders with distinct arrest trajectories even on the high-end of the criminal propensity continuum where the serious youthful offenders are located.

H_2: There are more than two groups of offenders with distinctly different arrest trajectories even on the high-end of the criminal propensity continuum.

H_3: There is an adolescence-peaked group even in samples of serious youthful offenders.

H_4: The age-crime curve is invariant among the latent classes of serious youthful offenders. Between-group differences will not vary across time.

These hypotheses are largely based on both the prior empirical results from the Laub et al. (1998) and Piquero et al. (2001) studies that indicate there is a significant level of heterogeneity in the offending patterns of serious youthful offenders, as well as the theoretical arguments of Cohen and Vila (1996) and D'Unger et al. (1998) that hypothesize the possibility of greater heterogeneity on the far end of this continuum than

previously suggested. H_2 and H_3 are central to Moffitt's (1993, 1997) theoretical perspective, while H_4 is central to the theories of Gottfredson and Hirschi (1990) and Sampson and Laub (1993). Evidence supporting H_2 would cast doubt on the adequacy of Moffitt's theory that there are *only* two offender groups in the population. Evidence supporting H_4 would support Gottfredson and Hirschi (1990), while evidence refuting it would support the theoretical positions of Sampson and Laub (1993).

Findings in support of these hypotheses are important the literature because serious youthful offenders are often referred to as being "relatively homogenous" (Ge et al. 2001: 750). As a whole, serious youthful offenders are an elusive class of offenders because they are (fortunately) relatively rare in the population of offenders (Cernkovich et al. 1985). Researchers thus are often forced to empirically "lump" together offenders who have met some minimum definitional criteria that usually involves a measure of either seriousness and/or chronicity of offending (McDermott 1983). After "making the cut," this group of offenders is usually isolated and treated as a homogeneous group (often labeled as the "chronic offender" group).[12] If there is significant heterogeneity in the propensity to offend within this population, recognition of that fact is important to the crime literature for both its theoretical and public policy implications.[13]

---

[12] Loeber et al. (1998) provide an extensive discussion of the variable cut-off points that have been used in an attempt to isolate the type of offenders used in this study.

[13] For example, in the article, "The Development of Persistent Criminal Offending in Males," Ge et al. (2001) analyze the arrest patterns of a sample 2,263 male committed to the CYA in 1964 and 1965. The authors analyze the arrest patterns of the CYA wards at ages 18-20, 21-25, 26-30, and over 31 using a series of ordinary least squares regression models. The authors' conclude (from a state dependence position) that, "early problem behaviors exert a significant influence on persistent offending. Early involvement with alcohol and drug use was a significant predictor of adult arrest frequency. This suggests that early substance use and abuse can influence criminal behavior throughout the life span." No attempt was made to control for unobserved heterogeneity, however, and thus it could simply be (and as would likely be posited by Gottfredson and Hirschi) that early drug use and abuse is correlated with the unmeasured (or at least the poorly measured) heterogeneity in the propensity to offend. Without controls

The results concerning these four hypotheses will be presented in Chapter 7. In that chapter, we apply Nagin and Land's (1993) semiparametric mixed Poisson model to each of the three samples used in this study. After determining the optimal number of latent classes present in each sample, the offending trajectories will be graphed over the age distribution. Comparisons of the trajectories will be made concerning the patterns of offending displayed over time within and between the latent classes.

STUDIES OF THE RELATIONSHIP OF PAST TO SUBSEQUENT CRIMINAL ACTIVITY

We now turn our attention to reviewing previous studies of the second critical issue addressed in this study—the relationship between past and subsequent criminal activity. Since investigating the relationship between past and subsequent criminal activity requires longitudinal (panel) data on a set of individuals, the studies reviewed here are limited to those studies following a panel of individuals over time. Furthermore, given the differential explanations of the population heterogeneity and state dependence positions for the underlying causes of the correlation between criminal activity at two different points in time, all of the studies reviewed here also control for individual differences in the propensity to commit criminal acts.

Historically, controlling for differences in criminal propensity has been most often attempted by including control variables measuring individual characteristics or other factors considered relevant in a regression model. However, multiple studies

---

for unmeasured heterogeneity, this finding is subject to serious criticism as a mere methodological artifact that would disappear in a more appropriate statistical model. See the section below describing studies of the relationship between past and subsequent criminal activity for a further discussion of the importance of unmeasured heterogeneity

(Paternoster and Brame 1997; Bushway et al. 1999; Nagin and Paternoster 2000; Paternoster et al. 2001) argue that there are two principal problems with this method of controlling for individual differences. First, criminologists cannot agree on the precise and most appropriate measures that reasonably capture individual differences in criminal propensity. Second, even if there was such a consensus on relevant variables, most data sets probably would not have some, most, or perhaps any of those key measures, making the task of adequate measurement impossible.

The end result of such problems, as noted by Nagin and Paternoster (2000: 131), is that "researchers would have no way in knowing if they have captured a sizeable share of the between-individual variation in criminal propensity with the measures they have available. Consequently, perhaps the lion's share of criminal propensity would be unmeasured or unobserved." Simulations by Bushway et al. (1999) show that failing to account for unobserved heterogeneity leads to seriously biased estimates that favor the state dependence argument (see also Heckman 1981a; Hsiao 1986). Unobserved heterogeneity is, in essence, akin to omitting a key variable from the model specification, resulting in biased estimates of the other included covariates that are correlated with the omitted variable (Bushway et al. 1999; Nagin and Paternoster 2000). Since prior criminal activity will be positively correlated with criminal propensity, failure to adequately control for persistent unobserved heterogeneity (in criminal propensity among individuals) will lead to positively biased coefficients for the variable representing prior criminal activity. In other words, without controlling for persistent individual differences, the coefficient for the prior criminal activity variable will absorb the effect of the omitted variable (individual differences in criminal propensity), resulting in an

overestimate of the effect of prior criminal activity on present criminal activity. As Nagin and Paternoster (1991: 10) explicate, "the problem is that association between the response variable and some specific covariate at the level of the individual is confounded with variation in the persistent unobserved heterogeneity across the population."

Thus, all of the studies to be reviewed below make use of statistical techniques controlling for unobserved or "hidden" heterogeneity. These studies have used one of two methods (and in one case both methods) to account for unobserved heterogeneity: (1) parametric random effects models or (2) semiparametric random effects models.[14] The primary difference between the two methods concerns the distribution of the unobserved heterogeneity. The parametric random effects models assume that the unobserved heterogeneity is *continuously distributed* in the population according to some known (mathematically tractable) parametric distributional form (e.g., it is normally distributed). The semiparametric form of the models *nonparametrically approximate* the form of the unobserved heterogeneity, assuming only that an approximation can be accomplished using a discrete, multinomial distribution (Heckman and Singer 1984; Nagin and Land 1993; Land et al. 1996). This semi-parametric random effects model is, in fact, the finite mixture model of Nagin and Land (1993) previously discussed. Each "latent class" is

[14] For readers desiring more information on these models at this point, these models are described in greater detail in Chapter 5. Briefly, the parametric random effects model assumes that the error term for an individual at any point in time is composed of two components: a time-invariant, individual-specific term and a pure random disturbance term (that is distributed according to some parametric assumption, usually the normal distribution). The individual-specific component, which is invariant over time, captures persistent, unmeasured individual differences in the propensity to offend. The correlation of an individual's error term over time (referred to as rho or ρ) is calculated as the variance of the individual-specific terms divided by the variance of the total error term (Hsiao 1986; Nagin and Paternoster 1991). Rho (ρ) is an estimate of the proportion of the variance of the error term that is due to persistent (time-stable) heterogeneity. If $\rho = 1$, the variance of the error term is entirely due to heterogeneity, whereas if $\rho = 0$, then persistent heterogeneity is negligible (Nagin and Paternoster 1991).

assumed to be a single "point-of-support" or "segment" of the multinomial distribution, and the distribution of unobserved heterogeneity (known as the mixing distribution in statistics) is approximated using a finite number of points-of-support. Within each "segment" of the sample, individuals are internally homogeneous with respect to criminal propensity, but individuals from different segments have varying propensities to engage in criminal activities.

As Bushway et al. (1999) note, both of these models make assumptions, and the degree to which the assumptions are tenable is key to the robustness of any observed results. The parametric form of the models is more restrictive and more efficient than is the semiparametric form, which is less restrictive and hence also less efficient. Violations of the assumptions of each model, including the assumption of the distribution of unobserved heterogeneity, can have a significant impact on the conclusions based upon the results obtained from each model. We will return to this significant issue of "violating assumptions" later in our discussion of the studies that address the relationship between past and subsequent criminal activity.

First, however, the results of several studies will be reviewed as they were reported in the original articles. In the following discussion, we try to stay substantively focused, but will include methodologically technical comments and footnotes when necessary. It should be stated that modeling the relationship between past and subsequent criminal activity is methodologically complicated, a point that should not be underemphasized (Nagin and Paternoster 2000). The old adage, "the devil is in the details," is quite appropriate for this issue. Table 3.2 presents key information from the

Table 3.2 Summary Information of Prior Studies of the Relationship Between Prior and Future Criminal Activity

Authors	Data Source	Sample Size	Gender	Periods Studied	Risk Level[A]	Dependent Variable	Finding[B]
Nagin and Paternoster (1991)	1981 South Carolina High School Survey	1163	Males/ Females	2	Low	Binary Indicator of Self-Reported Property Crimes	SD
Nagin and Farrington (1992)	1961 Cambridge Study	403	Males	11	Low	Two-year Binary Indicator of Convictions	Mixed
Sampson and Laub (1993)	1950 Glueck Study	690	Males	2	High	Arrest Counts Between Ages 25-32 and 32-45	Mixed
Laub and Nagin (1996)	1961 Cambridge Study	403	Males	11	Low	Two-year Binary Indicator of Convictions	Mixed
Paternoster et al. (1997)	1988-1989 NC Div. of Youth Services Releasees	838	Males	4 - 6	High	Yearly Arrest Counts	Mixed
Paternoster and Brame (1997)	1977 National Youth Survey	479	Males/ Females	5	Low	Binary and Count Indicators of Self-Reported Delinquency	Mixed
Bushway et al. (1999)	1958 Philadephia Birth Cohort	13160	Males	7	Low	3-Year Binary Indicator of Police Contact	Mixed

[A] Risk level is defined here as "Low" and "High." Low-risk samples correspond to general population samples that are likely to include a majority of low-risk cases in the data. High-risk samples, on the other hand, is used to refer to samples where high-risk cases will constitute the majority of cases (e.g., samples of parolees).

[B] Finding: SD = State Dependence; PH = Population Heterogeneity; Mixed = Both State Dependence and Population Heterogeneity.

different studies to be reviewed below that examine the relationship between past and subsequent criminal activity.

One of the key studies regarding the relationship between past and subsequent criminal activity was that of Nagin and Paternoster (1991). It is a study of 1163 respondents in a convenience sample of students from nine high schools in South Carolina. This panel began in 1981 and had subsequent "follow-up waves" in 1982 and 1983. Although there were initially 2700 sophomore respondents, at the final wave only 1250 senior respondents remained in the study, and of these, only 1163 filled out the information on relevant covariates deemed necessary for inclusion in the study by the authors. The dependent variable was constructed as a binary variable representing self-reported participation in three types of property crimes: stealing something valued less than $10, stealing something valued between $10 and $50, and breaking into a building and stealing something. Most respondents who indicated they had participated in one of these crimes had only stolen something valued less than $10. The panel assessed the dependent variable at two different points in time: participation between waves 1 and 2, and participation between waves 2 and 3. As the authors point out, the length of this panel (2 points in time) is the absolute minimum number of periods needed to estimate a panel model. In their model, the lagged dependent variable (participation in crime during the prior measured period) is the parameter estimate providing evidence for or against a process of state dependence.

Using a (parametric) random effects probit model, Nagin and Paternoster found a significant correlation between participation in the property crimes at the two points in time, even after accounting for unobserved heterogeneity through the use of the random

effects model. Participating in the property crime between waves 1 and 2 significantly increased the odds of participating in crime between waves 2 and 3, net of persistent heterogeneity (rho was estimated to be equal to 0 indicating that persistent heterogeneity was negligible). According to Nagin and Paternoster (1991: 183), "the results revealed that prior participation in crime had a positive and significant association with subsequent participation, controlling for the possibility of unobserved heterogeneity. This finding is consistent with the hypothesis that prior participation reduces the barrier to subsequent participation in crime. The authors were notably cautious in their conclusions, noting that it was an "exploratory" study with "suggestive results," because of several methodological limitations including the use of a non-random convenience sample, the built-in assumptions of the random effects model concerning the distribution of the heterogeneity (i.e., that it was normally distributed), and because the "initial conditions" assumption of the model was clearly violated.[15] Furthermore, due to heavy sample attrition (57% dropped out before the final wave), the potential "homogenization" of the sample with respect to criminal propensity could not be ruled out.

In addition to presenting some initial findings on the relationship of past to subsequent criminal activity, the study by Nagin and Paternoster (1991) was also noteworthy because (1) it was the first study to explicitly address and elaborate the state dependence versus population heterogeneity arguments for crime, (2) they proposed the use of the random effects models as a viable method for addressing the issues

[15] The initial conditions assumption refers to the assumption that at the first wave of the study, none of the respondents had already initiated the process (i.e., been involved in property crime activity). This assumption is required so that the model is able to obtain an unbiased estimate of the individual-specific component of the error term. It turns out that this assumption is utterly critical to calculating unbiased estimates concerning the relationship between past and subsequent criminal activity (Brame et al. 1999). See the discussion section below and Chapter 5 for further information.

surrounding the continuity in offending patterns, and, perhaps more importantly, because (3) their findings were so provocative as to stimulate continued research on the issue.

Soon after the publication of the Nagin and Paternoster (1991) study, Nagin and Farrington (1992a) presented results bearing on this issue using the data from the Cambridge Study in Delinquent Development previously described. Recall that this study employs 22 years worth of conviction data covering ages of 10-32 for 403 males, based on a model with 11 time periods. The dependent variable for each period was a binary indicator of conviction during consecutive two-year periods (e.g., any conviction during ages 10 and 11 constituted the offense or dependent variable for the first period). Following the lead of Nagin and Paternoster (1991), Nagin and Farrington (1992b) also used the random effects probit model that assumes unobserved heterogeneity to be normally distributed. Interest focused on the parameter estimate for the binary variable that indicated whether or not the individual had been convicted in the prior period (i.e., this is the lagged dependent variable). The coefficient for that variable represents the estimate of the state dependence effect for these data.

In contrast to the findings of Nagin and Paternoster (1991), this study found a highly significant, strong effect of persistent unobserved heterogeneity that served to significantly reduce the association between past and subsequent criminal activity. In the model ignoring persistent heterogeneity (i.e., a standard probit model), the parameter estimate relating conviction in the prior period to conviction in the subsequent measured period was 1.16. In the model controlling for persistent heterogeneity, the estimate was reduced in magnitude to 0.446, roughly a 62% reduction in the magnitude of the effect. Rho (the within-individual correlation of the error term across time) was estimated to be

0.4, indicating that 40% of the unexplained error variance was estimated to be due to persistent hidden heterogeneity. Nagin and Farrington (1992a: 253) focused their attention on the reduction of the magnitude of the state dependence parameter after controlling for unobserved heterogeneity and the large magnitude of the rho estimate. They concluded that the results were most consistent with the population heterogeneity position and that "...evidence of true state dependence is limited. After controlling for persistent unobserved heterogeneity, the association between past and subsequent participation is greatly diminished." However, closer examination of their results more clearly supports a "mixed" model where both population heterogeneity and state dependence processes are at work (Paternoster et al. 1997). Nonetheless, this study was important because it demonstrated that population heterogeneity, if left uncontrolled, could have serious effects on the estimates of variables indicating evidence in support of the state dependence process.

Although not explicitly addressing the relationship between past and subsequent criminal activity, further analyses by Nagin and Farrington (1992b) of the Cambridge data also revealed strong effects of persistent unobserved heterogeneity. Employing the same data and statistical models from the previous study, Nagin and Farrington (1992b) investigated whether age of onset had a significant effect on the probability of conviction in the 11-period panel, net of the effects of persistent unobserved heterogeneity. The state dependence interpretation of the age of onset variable is that early conviction has a causal impact on the probability of subsequent criminal activity (i.e., conviction causes changes in their life circumstances, such as increasing the likelihood of association with delinquent peers or reducing one's social bond, that makes continuing in a life of crime

more likely), whereas the population heterogeneity interpretation is that the age of onset variable is merely a proxy measure indicating the level of criminal propensity (i.e., individuals with an earlier age of onset have the highest criminal propensity levels). In the model without controls for persistent unobserved heterogeneity, the age of onset variable was found to have a large and highly significant negative effect, indicating that as the age at first conviction increased, the odds of a subsequent conviction decreased.[16] However, in the model controlling for unobserved heterogeneity (i.e., in the random effects probit model), the inverse association between age at first conviction and the odds of a subsequent conviction was reduced to insignificance and near zero in absolute magnitude. In other words, the inverse association between prior and subsequent criminal activity was entirely attributable to the effects of persistent unobserved heterogeneity (i.e., due to time-stable differences).

In a subsequent empirical test of their theory, Sampson and Laub (1993) presented a two-period panel analysis of arrests counts between ages 25-32 (period 1) and ages 32-45 (period 2). These analyses also employed the use of the Glueck data described earlier, plus results from the matched "control group." Here "nondelinquent" cases were matched (case-by-case) to their "delinquent" pairs on the basis of age, ethnicity (e.g., Irish, Italian, German, Jewish), and neighborhood (n = 289 for the delinquents; n = 401 for the matched control group).[17] Sampson and Laub utilized a

[16] The actual specification employed the use of two variables to capture the effect of age of onset. One variable indicated that if the individual had ever been convicted in prior period, while the other variable indicated the actual age of onset. The use of two variables allowed the state dependence effect to decrease as the age of onset increased (i.e., it allowed the positive impact of the first variable to be magnified by an early age of onset)

[17] Sampson and Laub (1993) present sample sizes for the "pooled" datasets. Since their analyses are based on a two-period panel model, we have divided the pooled sample sizes by two to arrive at the sample size

generalized least squares (GLS) random effects model (which is in essence a random effects OLS model) and they found results consistent with those of Nagin and Farrington (1992a).

In the models estimated on both the delinquents and the control group, Sampson and Laub found significant levels of persistent unobserved heterogeneity (rho = 0.22 and 0.29 for the delinquents and control groups, respectively). Furthermore, they found that even after controlling for persistent unobserved heterogeneity, both the unofficial and official juvenile delinquency behavioral variables were positively and significantly related to observed crime frequencies between ages 25 and 45. These results were found in the models for both the delinquent group and the control group. Similarly, the results of their analyses also indicated that several other variables representing the state dependence position were significantly related to engaging in crime during adulthood even after controlling for persistent individual differences. For example, the models for both the "experimental and control" groups indicated that job stability had significant negative effects on adult crime frequency. This suggested that individuals with less stable job histories were more likely to engage in crime during adulthood, net of the effects of persistent individual differences and measures of juvenile offending frequency. According to Sampson and Laub (1993: 198-199), "these findings support the idea that state dependence underlies the effects of both prior crime and weak social bonds." When all the evidence presented by Sampson and Laub (1993) is viewed in total, however, their

data clearly support the "mixed" position—both population heterogeneity and state dependence processes were found to be present in the Glueck data.[18]

In a re-analysis of the Cambridge data used in their initial article, Land and Nagin (1996) present further evidence concerning the link between having a prior conviction (at any point in the individual's past) and the probability of a subsequent conviction at a given age. Employing their semiparametric finite (Poisson) mixture model, Land and Nagin (1996) find evidence to support the mixed position that dovetails squarely with the conclusions of Nagin and Farrington (1992a), who had analyzed the same data with the parametric random effects probit model.[19] The analyses by Land and Nagin (1996) were the first to use the finite mixture models (allowing for a nonparametric specification of the distribution of unobserved heterogeneity) to address the question of whether past evidence of engaging in criminal activity has a significant effect on subsequent criminality after controlling for unobserved heterogeneity. Consistent with the conclusions of their initial article, Land and Nagin (1996) found significant differences in

[18] Sampson and Laub (1996) analyzed the military arrest history of the samples and also found that controlling for "military fitness" (fitness for military service), prior delinquency had a positive effect on the number of arrests acquired during their military service. This effect was not significant though (and there was no control for persistent unobserved heterogeneity). In support of their theory, however, they did find that early entry into the military significantly improved the lives of the structurally disadvantaged and delinquent men. The military was a "turning point" in the lives of these men, allowing them to better their lives, including their occupational status, job stability, and socioeconomic achievement in adulthood. This effect was especially pronounced among the veterans previously stigmatized as an official delinquent (i.e., processed through the juvenile court). In other words, events during adulthood have important consequences for subsequent outcomes. It should be noted, however, that Laub and Sampson (1998) also present evidence in support of the population heterogeneity argument. In these analyses, Laub and Sampson (1998) found that the delinquent group was significantly less likely to take advantage of educational opportunities available both while in the military and through the G.I. educational bill, and that chronic offenders were significantly less likely to enter into good marriages.

[19] Technically speaking, Land and Nagin (1996) estimated a multiple-spell discrete-time hazard model of the time until conviction (i.e., years until or years between convictions). Land and Nagin (1996) show that under regular conditions, the microlevel Poisson model is equivalent to a discrete-time hazard model. See Land et al. (2001) for further information regarding this event history formulation of the finite mixture model.

the propensity to offend within the sample. They uncovered the same four distinct trajectory patterns reported in their earlier article. There were significant between-group differences, however, in their longitudinal offending trajectories. The probability of surviving without a conviction to a given age varied dramatically across the groups. For example, at age 16 the survival probabilities for the "high-level chronics," "adolescent-limiteds," and "low-level-chronics" were 0.295, 0.532, and 0.800, respectively, while the corresponding survival probabilities at age 30 were 0.049, 0.443, and 0.445.

Land and Nagin also included a variable in their models to assess whether or not the individual had ever been previously convicted. The parameter estimate for the prior conviction variable indicated that individuals with a prior conviction had a higher "hazard" of being convicted at a given age. Notably this parameter estimate was calculated net of the effects of unobserved individual differences that were captured through the use of the points-of-support approach to unmeasured heterogeneity in the offending population. Stated differently, if you compared an "unconvicted" individual with an already "convicted" individual *within* the same "segment" or "point-of-support" of the unmeasured heterogeneity distribution and at the same age, the individual who had been convicted at a prior age had a much greater chance of being convicted at that age. For example, at age 16 the probability of "onset" (or first conviction) *within* the "high-level chronics" group was 0.321, whereas the "post-onset" probability of conviction *within* this group was 0.695.

It is worth noting that in their initial article, Nagin and Land (1993) also modeled a state dependence variable (lagged indicator of conviction in the prior period), but they did so within the "intermittency" portion of their model that only included controls for

observed heterogeneity (through the inclusion of observed variables) rather than unobserved heterogeneity.[20] The intermittency concept allows for the possibility that periods of criminal activity may be interspersed with periods of inactivity, yet that this inactivity does not signal the end of an individual's "criminal career." This portion of the model substantively investigates the factors that predict the probability of being an "active" offender at a given age. Nagin and Land specified this component of the model to be predicted by age, age squared, a lagged indicator of conviction in the prior period, and a composite, time-stable measure called TOT (that was composed of measures of risk taking attitude, parental criminality, a poor parenting indictor, and IQ). The parameter estimate for the lagged conviction indicator was significantly and positively related to the probability of being an active offender at a given age. It is interesting to note that the parameter estimate (1.09) for the lagged conviction indicator was nearly identical to the parameter estimate that Nagin and Farrington (1992a) found (1.16) in their standard probit model which did not account for unobserved heterogeneity. The intermittency component of the Nagin and Land (1993) analyses was important since it explicitly demonstrated the presence of within-individual variation in criminal offending, but the authors noted that the theoretical importance of the concept of intermittency was problematic (see also Horney et al. 1995).

Subsequently, however, Horney et al. (1995) proposed an explanation for the periods of intermittent offending. The authors connected the possibility of periods of activity being interspersed with periods of inactivity by drawing on the theory of

[20] The technical aspect of the intermittency component of their model is that it allows their semiparametric mixed Poisson model to be generalized as a "zero inflated Poisson" model that allows for more zeroes than would be expected by the standard Poisson distribution (Mullahy 1986; Lambert 1992; Zorn 1998).

Sampson and Laub to argue that the "intermittency" effect could be explained by short-term changes in the "local life circumstances." Adopting a clear "mixed" position, they note (Horney et al. 1995: 658-659),

> Although a persistent underlying trait like self-control can influence both an individual's overall level of offending and his or her stability of marriage and employment, that shared influence does not mean that a relationship between offending and the life circumstances is necessarily spurious. It is still possible that involvement in those social institutions influences the likelihood of offending *during the time of involvement*. The high crime rate of the most persistent offender, rather than indicating a total lack of investment in social institutions, may instead reflect alternating periods of criminal activity and inactivity.

Using data on short-term variations in both social bonding variables (such as going to school, working, living with a wife, drinking heavily, using drugs) and short-term variations of offending behaviors among a sample of incarcerated prisoners, the results of their hierarchical linear models showed that short-term, within-individual changes in offending behavior were strongly related to changes in the local life circumstances of the offenders, net of controls for unmeasured heterogeneity in the propensity to offend. The men in this sample (600 serious adult offenders) were significantly less likely to be involved in criminal activity when they were working, were not using drugs or alcohol, and were living with their wives. This finding is entirely consistent with a "state dependence" position and clearly highlights that short-term change in the offenders criminal activity is intrinsically related to short-term improvement (or worsening) of their local life circumstances. The implication is that if criminal arrest/conviction "worsens" the local life circumstances of offenders through its "negative effects" on the odds of

obtaining a job, going to school, or living with a wife, then (even after controlling for criminal propensity) a strong association between previous and current offending is to be expected.

A more recent study by Paternoster et al. (1997) followed the example of Sampson and Laub (1993) by examining the offending patterns of a sample of high-risk youthful offenders. Using a sample of 838 young, male offenders released from the training schools of the North Carolina Division of Youth Services in 1988-1989, Paternoster et al. (1997) examined the yearly arrest counts on the offenders between the date of release and November 1994, when the follow-up period ended. Using the random effects negative binomial panel model, the results of their analyses are based on 4-6 years of arrest counts (i.e., an unbalanced panel) from the post-release period. The unobserved heterogeneity was assumed to be distributed according to the beta distribution.

Similar to Nagin and Farrington (1992a), Paternoster et al. (1997) present results from both the standard negative binomial model with and without random effects for unobserved heterogeneity. A comparison of the log-likelihoods from the two models was used to test for the presence of significant persistent unobserved heterogeneity in the data, and a comparison of the results led the authors to conclude that there was a highly significant level of unobserved heterogeneity present in the data. The link between past and subsequent criminal activity was ascertained through the parameter estimate for the variable indicating whether the individual had been arrested in the previous year. The parameter estimate for the "state dependence" variable was 0.631 and highly significant (t-value = 8.82) in the negative binomial model that only controlled for heterogeneity through the inclusion of observed (measured) covariates such as previous juvenile

adjudications, race, gender, child abuse, family structure variables, and parental criminality. The parameter estimate was reduced to 0.228 in the random effects model allowing for both measured and unmeasured heterogeneity, indicating a substantial reduction in the magnitude of this effect (64% reduction in absolute size) after controlling for persistent individual differences. However, it should be noted that there was still a significant and positive effect even after allowing for persistent individual differences in the proclivity to offend. Thus, these results also indicate support for the mixed position that allows for both population heterogeneity and state dependence processes as causes of criminal offending.

Paternoster et al. (1997) also tested for differential effects of the state dependence process between the life-course-persistent and the adolescent-limited offenders as hypothesized by the Moffitt (and by Patterson). Recall that the arguments of Moffitt (1993, 1997) led to the conclusion that the offending patterns of the life-course-persistent group should be dominated by a static, population heterogeneity process (that has run its full course by the end of childhood/beginning of adolescence), whereas the offending patterns of the adolescent-limited group should be dominated by the state dependence processes and should be relatively unaffected by variables representing individual differences.

Age at first adjudication was used by Paternoster et al. (1997) as a proxy variable representing whether the case is a life-course-persistent (high criminal propensity) or an adolescent-limited (low criminal propensity) offender. In a series of models (14 separate models to be exact), the authors use different cut-points for the age at which to divide the sample into the low and high criminal propensity groups and then test for differential

state dependence effects on the basis of models run on the low and high criminal propensity groups separately.[21] The authors found age 15 to be a cut-off point that generated different (arguably minor) state dependence effects between these two groups. They also found, however, that any estimated differential effects will be highly sensitive to the age (at first adjudication) used to divide the sample into the two groups. As noted by Paternoster et al. (1997: 261), "this result [the age 15 cut-off point] strikes us as being consistent with the predictions offered by developmental theorists, but the lack of robustness in this result to slight variations in the early/late onset sample division scheme leaves us with some question as to whether the result is artifactual."

In a different study using the panel data from the National Youth Survey, Paternoster and Brame (1997) investigated the relationship between past and subsequent criminal activity among a sample of more "conventional" youths than those studied in the earlier Paternoster et al. (1997) study. This study used the 479 respondents that were age 11 or 12 in the first wave of the NYS study and followed them over the next 4 waves of the study (to age 15 or 16). Then Paternoster and Brame estimated random effects probit and negative binomial models on the binary and count variables, respectively, reflecting self-reported delinquent activity. The delinquent acts used in constructing the measures were theft exceeding $5, motor vehicle theft, aggravated assault, sexual assault, gang fights, robbery and breaking and entering (e.g., burglary). Paternoster and Brame were interested in the effects of two state dependence variables: a binary variable indicating

[21] Specifically, they tested whether the difference in the effect of the indicator of arrest in the prior period was statistically different from zero between the two groups.

delinquent behavior in the prior period (wave) and exposure to delinquent peers (proportion of friends who engage in delinquent acts).

The results of their analyses indicated that both prior delinquent activity and exposure to delinquent peers were positively and significantly related to current criminal activity. Controlling for persistent unobserved heterogeneity, which was found to be highly significant, the parameter estimate of the indicator of criminal activity in the prior wave was 0.86 in the probit model (participation model) and 0.984 in the negative binomial model (frequency model). The estimate from the logit model was very similar in magnitude to the estimate of Nagin and Paternoster (1991) in their analysis of the offending patterns of another "conventional" sample. Similar to Paternoster et al. (1997), Paternoster and Brame also test whether separate, distinct models (as hypothesized in Moffitt's theory) are necessary for describing the offending patterns of the life-course-persistent (early starter) and adolescent-limited (late starter) offenders.[22] After dividing the sample into two groups on the basis of their "offending propensity" at age 12, separate models were estimated on the two groups. The findings of these analyses also indicate (like the Paternoster et al. study) that there is no strong "statistical evidence that

[22] One of the primary differences between the three theories examined in this study can be viewed in light of the complexity of each theory (Land 1992; Paternoster and Brame 1997; Paternoster et al. 1997). The general theories of Gottfredson and Hirschi and Sampson and Laub assert that there is a single causal explanation for criminal offending over the life course that applies to all individuals in the population. Gottfredson and Hirschi's theory posits that a single causal pathway, which has run its course by the end of childhood, is all that is necessary to explain all of the variation in offending patterns (e.g., onset, frequency, desistence). Sampson and Laub's theory, however, is less parsimonious than the theory of Gottfredson and Hirschi because it posits that the precise predictors of criminal behavior vary as a consequence of the varying sources of social control experienced across the life course. The typological theory of Moffitt is the most complex theory of the three because it relaxes the assumption of a single etiology of criminal behavior and posits that there are two distinct causal pathways to crime. Paternoster and colleagues (Paternoster et al. 1997; Paternoster and Brame 1997) differentiate the three theories on the basis of being a general theory (Gottfredson and Hirschi), a dynamic general theory (Sampson and Laub), and a pure developmental theory (Moffitt).

these dynamic variables [prior offending and delinquent peers] exert different effects within groups of youngsters defined on the basis of their offending propensity at ages of 11 or 12 years and followed well into adolescence" (Paternoster and Brame 1997: 74).

In a recent discussion and comparison of three different methodological approaches to studying the relationship between past and subsequent criminal activity, Bushway et al. (1999) have presented an empirical application of the three different statistical approaches (random effects, semiparametric random effects, and fixed effects models) with the 1958 Philadelphia Birth Cohort data.[23] Using the 13,160 males in the birth cohort and 7 periods of data covering ages 6-26 (each period covered three "age years"), Bushway et al. (1999) apply a parametric random effects probit model, a semiparametric random effects probit (like that used by Nagin and Land), and a conditional fixed effects logit model to a binary indicator of police contact during each period.[24] This was the first presentation and application of the conditional fixed effects model as a potential methodological approach to study the processes of continuity and change in criminal offending over time. Similar to the previous studies, the state

[23] Land et al. (1996) also modeled the longitudinal offending patterns of the 1958 Philadelphia Birth Cohort. Land et al. (1996) did include a state dependence variable (lagged indicator of conviction in the prior period) in their specification, but they did so within the "intermittency" portion of their model that only included controls for observed heterogeneity (through the inclusion of observed variables) rather than unobserved heterogeneity. Nonetheless, the parameter estimate for the state dependence variable (in the intermittency part of their model) was 0.907 and highly significant.

[24] Conceptually, the conditional fixed effects logit model controls for persistent (time-stable) unobserved heterogeneity through the inclusion of a separate "intercept" or constant for each individual, although for numerical reasons these "intercepts" are "conditioned" out of the likelihood function during estimation. In other words, this estimator makes no assumption about the mixing distribution. However, a significant limitation of the model is that it does not permit the use of exogenous variables including age or trend variables. As noted by Bushway et al. (1999), the strong age or "trend" effects associated with criminal activity makes this is a serious limitation of this model (see also Maddala 1987; Hamerle and Ronning 1995)

dependence effect was modeled though the inclusion of a binary variable indicating police contact in the prior time period.

Results from two random effects probit models were presented first. The first model did not control for age ("trend") effects, and resulted in a parameter estimate for the state dependence variable (police contact in prior time period) of 1.052, which was highly significant (t-value = 48.13). The second model, which did control for time trend effects, produced a numerical estimate of 0.611 (t-value = 25.23) for the lagged police contact variable, indicating that part of the state dependence effect was partially the result of general temporal shifts in the probability of police contact. Both of these random effect models produced highly significant estimates of persistent unobserved heterogeneity (rho = 0.120 and 0.331 in the first and second model, respectively).

Next, semiparametric probit models with two (no time trend controls) and three (time trend controls) points-of-support were applied to the data. In the model with no time trend controls, the state dependence effect estimate was 1.035 (t-value = 49.25), while in the model with time trend controls the estimate was 0.608 (t-value = 23.72). It is interesting to note the nearly identical estimates of the state dependence effects from both the parametric and semiparametric formulations of the probit model.

The results of the conditional fixed effects logit model, which specifically limits the independent variables in the model to the lagged dependent variable only, estimated the state dependence effect to be 1.591. After translating the "logit" coefficient into "probit units" by dividing the estimate by 1.6, the estimate was essentially identical (0.994) to the estimates from the parametric and semiparametric models with no trend controls.

Discussion and Hypotheses Related to the Past and Subsequent Crime Relationship

There are two clear themes in the past decade of research on the relationship between past and subsequent criminal activity. First, there is unquestionably a significant amount of population heterogeneity in the propensity to commit criminal acts. Population heterogeneity was found to be significant in the sample of "conventional" respondents (Paternoster and Brame 1997), samples that over-represent individuals from an urban area (Bushway et al. 1999) and also in a predominantly working class area (Nagin and Farrington 1992a; Land and Nagin 1996), as well as in samples consisting of "high-risk" youthful offenders (Sampson and Laub 1993; Paternoster et al. 1997). Only one study (Nagin and Paternoster 1991) failed to uncover a statistically significant amount of unobserved heterogeneity in their sample. Given the possibility that sample selection processes over time (i.e., selectivity ultimately influences who was left in the sample at later waves in a panel) reduced the heterogeneity in this sample, this finding should be viewed with some caution. Despite the findings of significant population heterogeneity in offending patterns over time, Nagin and Paternoster (2000) note that the challenging assertions of Gottfredson and Hirschi (1990) are critical to the field of criminology (both theoretically and empirically) because they forced the discipline to acknowledge the importance of early life events, especially those within the family, and to consider how those events may have enduring consequences for individual behavior over time.[25] The controversial and provocative theoretical arguments of Gottfredson and

[25] Cohen (1987) has made similar arguments about the theoretical importance of the Wilson and Herrnstein (1985) population heterogeneity theory

Hirschi (1990) helped move the theoretical and empirical "lenses" of criminologists away from being obstinately fixated on the adolescent years.

The second theme stresses the importance of state dependence processes in the lives of offenders. All of the studies reviewed here found statistically significant evidence in support of the state dependence position. That is, controlling for unobserved heterogeneity in the propensity to offend, previous criminal activity was still positively and significantly related to the probability or frequency of current offending. Thus, despite individual differences in the propensity to offend, changes in the lives of offenders have important influences on criminal activity. Furthermore, these changes are beyond the explanation of a pure population heterogeneity argument. From the state dependence perspective, prior involvement in crime exerts a real (causal) effect on subsequent criminality though its attenuating effects on the social bond, and the constraints it places on future legitimate opportunities (Sampson and Laub 1993), and /or its disruptive effects on "local life circumstances" (Horney et al. 1995). "The empirical evidence indicates that whatever one's initial risk of crime, things can get better and they can get worse" (Nagin and Paternoster 2000: 137). Recall that the state dependence process is a double-edged sword contributing to both continuity and change in criminal offending patterns over time (Nagin and Paternoster 2000). Nagin and Paternoster (2000) note that the theory of Sampson and Laub has been important to the field of criminology for bringing the relevance of state dependence processes back into the view of criminologists after a period of time when the trend in criminology was to "push the causes of crime further back in the life course" (Grasmick et al. 1993: 5). Sampson and Laub reminded the discipline that events transpiring after adolescence have potentially

serious and important consequences for maintaining or changing previous behavior patterns.

Clearly, however, the summary that best characterizes the current research to date is that both processes are operating; that is, the evidence supports the "mixed" model where state dependence and population heterogeneity processes are necessary to explain both continuity and change in criminal behavioral patterns over time. In our judgment, the best example of the "mixed" position is found in the two studies that compare the magnitude of the state dependence effects both prior to and after controlling for individual differences in the propensity to offend. The studies we speak of were conducted by Nagin and Farrington 1992a, and Paternoster et al. 1997. In the standard probit models (without a correction for unobserved individual differences), the magnitude of the estimates are 1.16 and 0.631 for the Nagin and Farrington (1992a) and Paternoster et al. (1997) studies, respectively; whereas in the parametric random effects models, the corresponding parameter estimates are 0.446 and 0.228. Yet, even in the face of a roughly 63% reduction in the size of the parameter estimates (after controlling for unmeasured individual differences), the state dependence variables in both studies still remained positive, significant, and substantively meaningful. Thus, just as the pure state dependence perspective must concede a significant amount of the link between past and subsequent criminal behavior is due to persistent individual differences in the proclivity to offend, the pure population heterogeneity perspective must concede that prior individual differences cannot explain all of the association between criminal activity at two different points in time.

Recently, though, Brame et al. (1999) and Bushway et al. (1999) have raised some concerns about the validity of these important substantive conclusions as possible artifactual flaws in prior research. These potential flaws surround the two main assumptions of the parametric random effects models regarding: (1) the distribution of the unobserved heterogeneity (i.e., the mixing distribution) and (2) the initial conditions assumption.

First, the problem with the assumption concerning the mixing distribution is that the correct statistical inferences concerning (dynamic) state dependence variables require the mixing distribution to be correctly specified (Bushway et al. 1999). As it stands currently, there is no agreed upon distribution assumed to correctly capture the distribution of criminal propensity in the population (Land and Nagin 1996), and, further, the nature of the distribution may be very different in low-risk samples compared to high-risk samples. It was such uncertainty regarding the actual mixing distribution in the population that led to Heckman and Singer's (1984) "point-of-support" approximation (subsequently generalized by Nagin and Land to mixtures of Poisson models for event count data) whereby the continuous distribution, *whatever its shape*, is approximated by a discrete, multinomial distribution. The failure to correctly specify the unobserved heterogeneity distribution may result in both biased estimates and/or inflated significance tests (Bushway et al. 1999; Heckman and Singer 1984; Land et al. 1996). For example, simulations by Bushway et al. (1999) showed that when the actual distribution of unobserved heterogeneity becomes more skewed relative to the assumed normal distribution, the bias in the state dependence parameters becomes larger, thereby unjustly favoring the state dependence explanation.

Second, the initial conditions assumption requires that the behavioral process under study (e.g., criminal offending here) be observed at the *initial* start of the process (Hsiao 1986). Under this assumption, the lagged value will be zero for *all* cases during the first period under study (precisely because the process has not started). This condition ensures the lack of correlation between the lagged value in the first period and the time-stable (individual-specific) component of the error term in the model. Hsiao (1986) shows that it is this initial conditions process that allows the error term to fully incorporate heterogeneity in individual differences, and if one can meet this assumption, then the effect of the lagged value on the current value will be consistent even if the lagged outcome value in subsequent periods is correlated with the persistent unobserved heterogeneity (Brame et al. 1999).

As shown by Heckman (1981a) and Hsiao (1986), the main problem with violating the initial conditions assumption is that the parameter estimate for the lagged values of the outcome variable will be upwardly biased (i.e., favoring the state dependence perspective). The simulations of Brame et al. (1999) provide further evidence that a failure to meet the initial conditions assumption upwardly biases the estimate of the lagged value (i.e., the state dependence effect). As Brame et al. (1999: 612) note, "the failure to observe initial conditions guarantees that the parameter estimates from the random-effects model will be biased and inconsistent." The upwardly biased estimate is a direct consequence of the confounding of prior offending with the unobserved heterogeneity, whereby the parameter estimate for the lagged value absorbs some of the variation that should be rightly attributed to the time-invariant (individual-

specific) component of the error term that represents population heterogeneity (Heckman 1981a).

Brame et al. (1999) reanalyzed the data from the Paternoster and Brame (1997) study to see if the violation of the initial conditions in that data led to any erroneous conclusions regarding the impact of the state dependence variables. Using Heckman's (1981a) approximation method developed to correct violations of the initial conditions assumption, they found a further reduction in the importance of the state dependence variables (lessening the impact of the delinquent peer exposure variable and a complete reduction to nonsignificance of the prior offending variable) after correcting for violations of the initial conditions. The authors' concluded that, "reported coefficient estimates for dynamic factors could be biased because of problems with initial conditions" (Brame et al. 1999: 636).

As a result of such analyses, a general doubt lingers in the field about the robustness and validity of the findings of previous studies: "...in the absence of clear knowledge about fidelity to model assumptions, researchers should adopt a healthy amount of skepticism in their observed findings" (Nagin and Paternoster 2000: 140). Supporting this notion (in a critical essay on the superfluous treatment of the assumptions of statistical models), Maltz (1994) has persuasively warned that criminologists must devote more attention to checking the assumptions of the statistical models they apply to crime data or risk possibly generating publishable yet erroneous/invalid results. Consider the two following points regarding the seven primary studies reviewed above. First, five of the studies we reviewed earlier relied *entirely* on the parametric random effects model and made no attempt (for obvious reasons of both data and software limitations) to check

the robustness of their findings to the assumed distribution of the unobserved heterogeneity. The two notable exceptions that take advantage of the semiparametric formulation of the random effects model are the studies of Land and Nagin (1996) and Bushway et al. (1999). While the Nagin and Farrington (1992a) study did make use of the same data set as did Land and Nagin (1996), the state dependence variable relating past to current participation in criminal activity was specified differently in the two studies. The results of the analyses by Bushway et al. (1999) yielded virtually identical numerical and substantive results in both the parametric and semiparametric formulations of the statistical models. To date then, the only study that has calculated the state dependence effect (with the same data and the same exact model specification) with both the parametric and semiparametric random effects models yielded identical results in both models. As Bushway et al. (1999: 53) state, however, "since there is no reason to believe, a priori, that the results of our substantive analyses are generalizable beyond the specific data set that we used, we think that multiple-method strategies for investigating questions such as the one addressed here…are necessary." Thus, the degree to which the assumptions of the statistical models yield any substantive differences in the conclusions requires testing with other data sets such as those to be used in this study.

Second, only two of the six studies that employ the parametric random effects model use data that unquestionably do not violate the initial conditions assumption. These are the studies of Nagin and Farrington 1992a and Bushway et al. 1999. In the other four studies, the offending process had already been initiated at the point in time each study began their panel, and thus there is a possibility that the estimates of the

variables representing the association between past and subsequent criminal activity were upwardly biased.

Indeed, the two studies using samples of high-risk respondents (Sampson and Laub 1993; Paternoster et al. 1997) have relied entirely on the parametric random effects model and began their panel studies at a point in time when all of the respondents had already begun their criminal offending. Given the findings of Brame et al. (1999) and Bushway et al. (1999) regarding the consequences of violating these two assumptions, the presence and magnitude of the state dependence effects obtained in these two studies are currently questionable. This point is fundamentally critical because Nagin and Farrington (1992a), Paternoster and Brame (1997), and Nagin and Paternoster (2000) have all hypothesized that the importance of state dependence variables may depend on the nature of the sample employed in one's study. This assertion is based on the fact that studies showing stronger support for the state dependence perspective tend to employ low-risk samples (and self-report data), whereas studies showing stronger support for the population heterogeneity perspective tend to employ higher risk samples and use official data such arrest or conviction data.[26] Paternoster and Brame (1997) speculated that such findings are consistent with the theoretical propositions of Moffitt's (1993) dual taxonomy theory because samples consisting of higher risk cases should contain a larger percentage of life-course-persistent offenders (who's behavior is governed by static processes), whereas adolescent-limited offenders (who's behavior is governed by

[26] Again, it is important to bear in mind that while nearly all studies find evidence in support of both positions; it is the degree of support for each position that this issue concerns.

dynamic processes) should constitute the majority of the respondents in low-risk samples (see also Cernkovich and Giordano 2001).

In sum, there currently is some ambiguity regarding the state of the extant empirical evidence regarding the relationship between past and subsequent criminal activity, especially with respect to how robust the findings are to assumptions of the statistical models employed and how important state dependence processes are in the population of high-risk offenders. We concur with Nagin and Paternoster (2000: 131) that "only by examining the sensitivity or robustness of research findings with different statistical models and different data can the field hope to come to an understanding about the tenability of population heterogeneity and state dependence." Cernkovich and Giordano (2001) also have recently commented that there simply is "scant evidence" regarding the empirical importance of these two explanations (state dependence and population heterogeneity) for continuity and discontinuity of criminal offending patterns across the life course, especially in both the serious offender population and in data sets that include ages beyond adolescence.

In direct response to these calls for further investigations of this key theoretical issue, our study will test the following four hypotheses concerning the relationship between past and subsequent criminal offending behavior:

H_5: There will be a statistically significant positive association between past and subsequent offending behavior.

H_6: After controlling for persistent individual differences in criminal propensity, the association between past and subsequent offending will be reduced to a nonsignificant level.

H_7: After controlling for persistent individual differences in criminal propensity, the association between past and subsequent offending behavior will be reduced in magnitude but will still be positive and statistically significant.

H_8: The association between past and subsequent offending behavior will be nonsignificant for the life-course-persistent (high criminal propensity) group(s), while the effect should be substantial and significant for the "adolescent-peaked" group.

The key hypotheses for the three theories discussed in this study are H_6, H_7, and H_8. Evidence supporting H_6 would lend credence to the theory of Gottfredson and Hirschi, evidence supporting H_7 would be consistent with the predictions of Sampson and Laub's theory and evidence supporting H_8 would appear to validate Moffitt's theory.

Results concerning these four hypotheses are presented in Chapter 8 of this study. Here we will draw and build on the multi-method approach of Bushway et al. (1999) to test H_6 and H_7. More specifically, we will use both the parametric negative binomial random effects model and the semiparametric mixed Poisson model of Nagin and Land. In addition, we will also employ standard negative binomial models with a set of binary variables that indicate latent class membership (from the latent class results presented in

Chapter 7) to more definitively test the robustness of the presence (or absence) of state dependence effects in a longitudinal panel analysis of the offending patterns for three samples of California Youth Authority parolees.

To test the last hypothesis, H_8, separate models will be estimated on offenders assigned to a given latent class. It should be noted that H_8 is a conditional hypothesis that requires the identification of an "adolescent peaked" group in the data sets. To date, tests concerning H_8 have been accomplished after dividing the samples into two groups on more arbitrary grounds (e.g., age of onset) rather than calculating the effects within a group shown to actually offend in an adolescent-limited fashion. Before concluding the present chapter and moving on to Chapter 4, we discuss the potential contributions this study can make to the discipline of criminology.

CONTRIBUTIONS TO THE DISCIPLINE

The proposed research conducted herein will attempt to make several contributions to the discipline of criminology. These contributions include a general accretion of knowledge to the study of the continuity and discontinuity of criminal offending patterns over the life course of serious youthful offenders, as well as specific contributions that advance our current knowledge concerning both the relationships between age and crime and between past and subsequent criminal activity. A major contribution of this study centers on the nature of the samples employed in the analyses, the fact that three separate samples from different time periods are employed in the analyses, the length of time over which the subjects in the samples are followed, and the relatively large sample sizes are unique to this study.

Broadly speaking, the analyses to be presented should fill a void in the literature concerning crime over the life course and represent a foundational attempt at examining the long-term patterns of criminal offending among the most serious youthful offenders in the population. There are two primary reasons why we believe this to be so. First, the three samples analyzed here are relatively large, representative samples of youthful offenders who commit the most serious crimes at a disproportionately high rate.[27] This highly publicized group has to date been largely unavailable to social scientists (see e.g., Cernkovich, Giordano, and Pugh 1985; Cernkovich and Giordano 2001; Laub and Sampson 2001), and therefore a detailed empirical analysis of their longitudinal criminal offending patterns will provide social scientists and policy makers with a more accurate characterization and deeper understanding of the longitudinal patterns of criminal activity across the life course of this select group of offenders.

To date, much of our knowledge concerning the serious, persistent young offender has been derived through analyses of the most frequent offenders (usually referred to as the chronic offenders) in birth cohort studies such as the 1945 and 1958 Philadelphia birth cohorts and with general population samples. The major finding of these Philadelphia birth cohort studies was that about roughly six to seven percent of the individuals in the cohorts were responsible for more than half of all of the officially recorded police contacts reported for the cohort (see e.g., Tracy et al. 1985; Wolfgang et al. 1972). Wolfgang and his colleagues report in the 1972 study that these young "chronic criminals" were responsible for committing 63% of all known Index crimes

[27] The bulk of the analyses in this study involve the longitudinal offending patterns of 4,866 sample subjects (n = 1989 in the 1981-82 sample, n = 1443 in the 1986-87 sample, and n = 1434 in the 1991-92 sample)

committed by the birth cohort members (as measured through police contacts)--including 82% of all robberies, 73% of all rapes, 69% of all aggravated assaults, and 71% of the murders.[28]

However, several re-analyses of the 1945 birth cohort data question exactly how chronic or serious most of the chronic offenders in these types of samples (e.g., birth cohort studies or national probability samples) really are. For example, analyses by Bernard and Ritti (1991) indicated that only 35% of all police contacts in the Philadelphia birth cohort ever resulted in an actual formal arrest and that an astounding 31% of the 627 "chronic" offenders (those with five or more police contacts) were either never arrested (n = 48) or arrested only one time (n = 145) in their entire juvenile criminal career. Similarly, analyses by Weitekamp et al. (1995) showed that 73% of the aggravated assaults were committed by 32 of the 627 chronic offenders and 71% of the homicides were committed by only 10 of the chronic offenders. Studies of criminal offending that employ the use of general population samples generally contain too few serious youthful offenders (because of their low base-rate in the population) to allow for reliable descriptions and investigations of their offending behaviors over time (see e.g., Cernkovich, Giordano, and Pugh 1985; Loeber and Farrington 1998). For comparative purposes, it has been estimated that only 1 out of every 1000 police contacts will result in

[28] It was findings such as these that initially aroused interest in the application of selective incapacitation policies. The public policy implications of these birth cohort studies appear relatively straight forward to some criminologists and politicians: if the chronic offenders can be identified and segregated from the non-institutionalized population, then the rate of serious crime in a society could be reduced substantially (see, for example, Blumstein et al. 1986). Of course, selective incapacitation policies only work properly if the high-rate offenders commit crimes at a relatively constant, stable rate across age (Gottfredson and Hirschi 1986; Haapanen 1990; Ezell and Cohen 1997).

a single case being committed to the California Youth Authority (Legislative Analysts Office 1995).

Second, the length of time the cases in the three samples are longitudinally tracked allows for rigorous testing of the extent and nature of the patterns of continuity among this population. Standard (between-individual) empirical assessments of continuity in offending behavior often use cross-sectional data and/or short-term panel data that preclude most studies from addressing the main questions investigated in this study (as they require extensive longitudinal data):

> Most criminological research consists of cross-sectional "snap-shots" or short-term panel studies of crime over the full life span. As a consequence, relatively little is known about desistence and, for that matter, the processes of persistent criminal behavior through the life course. Indeed, the characteristics that distinguish persistence in a life of crime from desistence within any group of high-risk offenders are generally unknown (Laub and Sampson 2001: 1).

This study will addresses this limitation in the empirical literature by following three samples of offenders from the date of their first arrest until June 30, 2000.[29]

Laub and Sampson (2001) also point out that our knowledge of the long-term offending patterns among serious offenders has been hampered not only by a lack of studies that longitudinally follow this group for extended periods of time, but also because of the disjunction between the juvenile and adult record systems. Crime data often suffers from the division of juvenile and adult criminal record-keeping systems, meaning a dearth of data bridging the juvenile and adult years and that many datasets are

[29] For the three release samples combined, the earliest year of birth was 1956 and the latest was 1978. The 25th percentile for the year of birth was 1963, the median was 1967, and the 75th percentile was 1971.

often blind to criminal activity on the other side of the juvenile-adult age boundary (Blumstein et al. 1986; Cernkovich and Giordano 2001; Laub and Sampson 2001). The research undertaken in this study is unique because it entails following three samples of youthful offenders from the date of their initial arrest, through the year(s) they were incarcerated in the California Youth Authority, and then into adulthood, and well into adulthood for some of them (age 43 was the oldest age). The first sample was released in fiscal year 1981-82 and followed-up into their late 30s to early 40s (depending on age at time of release from the CYA). The second sample was released in fiscal year 1986-87 and followed into their early to mid 30s. Finally, the third sample was released in fiscal year 1991-92 and followed into their mid to late 20s.[30] We believe that these data sets comprise the most comprehensive set of longitudinal data on the serious youthful offender population that have yet been gathered. The length of the follow-up period and the sophistication of the statistical models used in this study allows for a rigorous examination of criminal offending patterns over the life course for these three data sets..

The empirical fact is that we currently know very little about the offending patterns of very serious youthful offenders; we simply lack evidence regarding fundamental questions concerning the criminal offending patterns these offenders. In a recent book surveying the current state of empirical knowledge on serious and violent juvenile offenders, the editors concluded with a section entitled, "Developing a Research Agenda." Here, they noted that there are currently many "gaps" in the knowledge concerning the nature and development of the longitudinal criminal offending patterns of

[30] For the 1981-82, 1986-87, and 1991-92 samples, the average ages at the end of follow-up period (June 30, 2000) were 37, 33, and 27, respectively.

serious and violent juvenile offenders, including the validity of offender typologies (e.g., life-course-persistent and adolescent-limited offenders) in this population, the nature and extent of the adult criminal offending patterns and adult life experiences of this group of offenders, and which covariates predict a continued persistence in offending within this group of offenders (Loeber and Farrington 1998). Similarly, Laub and Sampson (2001: 10) call for a theoretical and empirical focus on the patterns of continuity (persistence) and discontinuity (desistence) among samples of persistent and serious offenders, noting that "criminologists should not spend much time or energy studying termination and desistence for low-rate offenders" because such offending is normative during adolescence. The analyses completed in this study hence should provide needed evidence concerning the nature and extent of criminal offending across the life courses of the most serious youthful offenders.

Our review of the studies focusing on the relationship between age and crime as exhibited through the age-crime trajectories of discrete offender groups has indicated that there are several limitations with these analyses to date. The research presented in this study can contribute to the extant literature on this topic by addressing several of these current limitations. As noted above, previous empirical studies have often employed the use of data sets covering limited segments of the age distribution that preclude the formation of conclusions regarding the age-crime curve over the earlier or later ages. Researchers generally have only had access to a single sample which precludes addressing how stable or instable the various latent classes are across time. Finally, analyses of the age-crime curve within the population of very serious youthful offenders

have been extremely rare, especially over an extended portion of the age distribution and with contemporary samples.

In sum, we currently have very limited information concerning the actual shape of the aggregate age-crime curve within the serious offender population, on what "latent classes" of offenders are present in such a population, and if the identified latent classes are resilient across time. As Cernkovich and Giordano (2001: 405-406) note,

> if we accept the premise that there is a small group of offenders who do not begin and age out of crime in the same fashion as most offenders, then it is important that researchers examine in detail the extent to which the stability-change paradox is a function of the existence of two distinct population of offenders...however this issue has not been systematically examined, in part because of the relatively scant (though increasing) research focusing on serious chronic offenders...it is essential that the research agenda be expanded to include an even greater focus on this group and...that it include attention to patterns of antisocial activity prior to and beyond the adolescent years.

Furthermore, even though there is a group of offenders who have been semantically labeled "life-course-persistent," to date there is no convincing empirical evidence to prove that this group of offenders exists in sufficient numbers deserving of such a demonstrative label. A key question yet to be answered is: how persistent are the life-course-persistent offenders? Due to the highly selective nature of who gets committed to the California Youth Authority, samples of offenders released from the CYA have a unique potential to address this question. The data we describe in Chapter 6 aim to addresses these specific limitations by applying Nagin and Land's finite mixture model to

the three large samples of serious youthful offenders that are followed for an extensive period of time.

Our review of the prior studies of the relationship between past and subsequent criminal activity also made it clear that there is a critical need for the continued examination of this topic. As indicated above, previous studies have consistently found evidence in support of the mixed position; however, questions remain regarding the authenticity of the state dependence effects uncovered in the prior research, especially within high-risk samples, due to the possible methodological consequences of violating the assumptions of the parametric random effects models. The results to be presented in Chapter 8 of this study should contribute to the extant literature on this topic by examining the relationship between past and subsequent criminal offending.

Before getting to the methods and the data analysis chapters in this study, however, our attention in the next chapter is first focused on the California Youth Authority, the institution from which our samples have been released on parole.

CHAPTER 4

THE CALIFORNIA YOUTH AUTHORITY

INTRODUCTION

To successfully argue that these data represent California's most serious youthful offenders, readers must first understand the CYA's role in the state's judicial system and how offenders end up as "wards" of their institutions. Below we describe the mission and policies of the CYA, California's state agency responsible for housing, controlling, and rehabilitating the worst 5% of the youthful offenders in the state. Much of our description of the CYA focuses on the period between 1981-1992, the period most relevant for our offender samples.

THE CALIFORNIA YOUTH AUTHORITY

Historically, a disproportionate share of the attention of California's legal system has been directed at the affairs of its youthful offender population. For example, in 1992 about half of all persons arrested for crimes in the state were between the ages of 11 and 24, even though this age group consisted of only about 20% of the state's population (Legislative Analysts Office, 1994: 25). Between 1981 and 1992 (the period in which our sample members were released on parole), California's juvenile justice system would incarcerate a higher percentage of its youth for longer periods of time than any comparable state in the nation. Thus, at a time when many states were reducing their use of institutionalization, rates of incarceration for juveniles in California would be about twice that of the national average. Krisberg (1987), for example, reports that in 1985,

California's incarceration rate of youths ages 10-17 was 430 per 100,000 persons, compared to 125 in Texas, 126 in Illinois, 170 in Michigan and 230 in Ohio.

In California, adjudicated serious and repetitive youthful offenders are generally referred to the California Youth Authority (CYA). As we will see below, critics have charged that the state's decision to invest in large, isolated, and costly high-security facilities would produce an overcrowded system that offered expensive treatment that appeared to be of minimum value in rehabilitating wards. These same critics would argue that this crime control strategy failed to achieve its stated goal of securing the long-term protection of public safety for the residents of the state (see DeMuro et al. 1988). Defenders of the Youth Authorities' strategy for controlling youthful crime would, on the other hand, counter these criticisms with the argument that such high-security facilities were necessary because of the high propensity for violence among the volatile young men and women in its institutions (CYA, 1988). They argued further that the willingness of inmates to inflict harm on one another was learned on the streets and continued in the institutions. This behavior necessitated secure facilities. CYA supporters would therefore argue that the Youth Authority housed a dangerous population and this fact existed without regard to its crowded conditions. Below we briefly trace the history of some of the policies and programs at the Youth Authority and supply the reader with some necessary background on its development from its inception in 1941 through 1992, the last year of our sample period.

The Youth Authority's Origin

A separate juvenile justice system to deal with California's youthful offender population was first established in 1903. Until 1941, however, this system was largely a diffuse, county-by-county operation that lacked both integration and consistency. The California Youth Authority was initially established by the California Legislature after lawmakers became disenchanted with certain controversial activities reported to have occurred at the state's three existing juvenile correctional schools. Responding to a series of well-publicized escapes, reported scandals, suicides, and allegations of several forms of abusive treatment at these facilities, CYA legislation was officially signed into law by Governor Culbert Olson on July 9, 1941 and became effective as of September 13th of the same year. The Youth Authority was specifically established to serve as a sentencing alternative for young adult criminals and as a dispositional alternative for juveniles who had committed both criminal and status offenses. A main objective of this legislation was to produce a state-driven, integrated system whose major goal was the prevention of future illegal activities by youths that had come to the attention of the courts. For example, the statutory statement that perhaps best reflects the original intent of the legislation establishing the Youth Authority reads as follows:

> The mission of the Youth Authority is to protect the public from criminal activity by providing education, training, and treatment services for youthful offenders committed by the courts, assisting local justice agencies with their efforts to control crime and delinquency, and encouraging the development of state and local programs to prevent crime and delinquency (State of California, 1941:ch. 937).

The model that the legislators chose to follow was adapted from recommendations proposed in a report issued by The American Law Institute (ALI) in 1940. In this report

the ALI recommended the creation of a separate Youth Correction Authority in each of the United States.

The American Law Institute's Model Act

The American Law Institute became interested in redesigning the administration of criminal justice for the nation's youth in 1938 after concern was raised by its members over a long series of scholarly and media investigative reports that detailed the failings of the nation's juvenile justice system. Such reports had already made reform of the country's juvenile justice system a major theme among many civic and political interest groups worried about the health and welfare of the nation's young law violators. After conducting a survey of the nation's juvenile justice system in 1938, the ALI found this system to be "uncoordinated, inefficient, and ineffective." The ALI's national survey indicated that the juvenile justice system was conceived mainly for the purposes of punishment and not for rehabilitation, and that the quest for justice for young Americans was thus being severely thwarted. The institute soon organized a panel consisting of experts from various academic and professional fields to study the problem, and to formulate recommendations for an improved system of justice for troubled youths.

Thus in 1938, the Executive Committee of the ALI charged its panel with the task of developing a Model Act to improve the treatment of the country's juveniles who had been adjudicated for illegal offenses. After an extensive investigation, the panel's final draft was submitted to, and adopted by, the ALI in 1940. It was released to the public the same year under the title of the "Youth Authority Correction Act." This Model Act was to represent a guideline for states to follow. It provided for the authority to set up and

operate new institutions to deal with problem youth. The Model Act specifically called for the creation of a "Youth Correction Authority" in each state composed of three persons who were to oversee the employment of educators, psychiatrists, physicians, and social scientists, etc., who would assist the state in achieving the goal of providing for the correction and rehabilitation of youthful offenders. In effect, the Model Act called for the removal of all power from judges to determine the type and length of treatment to be accorded adjudicated youthful offenders. The only exceptions to this rule would be cases where the judge imposed a sentence of death, life imprisonment, or merely imposed a fine. The judge would, in the vast majority of all cases then, be limited to committing the youth to an indeterminate sentence at the Youth Authority.

The Model Act specified that all adjudicated youth who were not given a death sentence, life imprisonment, or a simple fine were to be bound over to the Youth Authority of each state for a through diagnostic evaluation. It would be the duty of the Youth Authority to give careful examinations to those who had been committed in order to determine the best treatment available to fit their individualized needs. It would also be the duty of the Youth Authority to see that such treatment and control were maintained until it was safe to return the youth back to the community. The Youth Authority, when prescribing treatment, would be permitted to utilize any existing public institution and agency within the state that they perceived to have the means to treat the youth's diagnosed condition. Thus, the Youth Authority could commit a youth to any state reformatory, parole or probation agency, etc., however, the agency would not be permitted to interfere with the operation of treatment given in these facilities. The Model Act also gave the Youth Authority the power to remove any youths from such facilities if

it believed these persons were not getting proper beneficial treatment. The Youth Authority would also be permitted to make use of any available private institutions or agencies or services within the state that it deemed suitable to best treat the youth's specific problems, if these facilities consented to such an arrangement.

The Model Act specified further that the Youth Authority, within economic limits, should be given the power to request appropriations from the state to create additional facilities if there were none available in the state in order to properly and adequately treat youths with specific types of problems. The key element of the Model Act then was to create a system that would allow for the appropriate diagnosis of each individual's needs, one that would have the flexibility to properly treat each individual offender after a clinical diagnosis was made.

The ALI emphasized in the preface of its final report that "not until the theory of the punitive system is discarded in favor of a corrective and preventive plan will repetitious crime be effectively checked (1940: xii)." In effect, the ALI's response to the problem of the existing punitive juvenile justice system was to advocate the creation of a single central administrative agency within each state with jurisdiction over all post-conviction procedures. The main goal of this agency would be to protect society through rehabilitation, not punishment. To accomplish such a purpose, the ALI model mandated change from an orientation to punish toward one that stressed the "rehabilitative ideal." The ALI report argued that this change could be best accomplished through the application of organizational theory and the social science research to the youth crime problem (Bolen, 1972). Thus the ALI was unique for the time period when advocating

the application of an organizational approach to solve the problems of youthful offenders in society.

In sum, the structure of the centralized agency recommended by the ALI was to have great flexibility to deal scientifically and intelligently with problem youth. Such youths were to be subjected to scientific diagnostic procedures that would determine their individual rehabilitative needs. Once the determination of their specific individual needs was made, the youth was to be referred to whatever existing service or facility in the state the Youth Authority considered most likely to benefit that person. Thus, in its final report, the ALI called for the establishment of a statewide organization that would promote deterrence through rehabilitation. This promise of rehabilitation was most likely to be obtained through the use of principles from organizational theory such as rationality and efficiency that could be effectively utilized to prescribe individualized beneficial treatment.

The ALI released its final report at the precise time that the California legislature was searching for a model that would quiet the controversy created by the aforementioned incidents at the state's three juvenile correctional schools. While California's Model Act followed closely the major principles outlined in the ALI's Model Act, there were some significant differences between the two models. First, California's Model Act increased the age over which the Youth Authority was to have jurisdiction over wards from 21 to 23 years, and the maximum age over which the state was to maintain control was increased from age 21 to 25. Also, while the ALI's Model Act dealt specifically with convicted offenders, California's Model Act authorized the Youth Authority to work towards delinquency prevention as well as rehabilitation. Finally, the

California Model Act attempted to limit political influence in the selection of members of the Youth Authority Board. It called for board appointments to be made by the Governor, but such appointments could be made only from a list of qualified persons to be selected by an independent advisory panel. The final selection of candidates from the list would then be ratified by the state legislature.

By establishing its Youth Authority in 1941, California would became the first state to officially endorse the American Law Institute's proposal for a central authority as the means to best coordinate and achieve the rehabilitative ideal through mechanisms informed by an administrative theory of corrections (Bolen, 1972: 3). As a result of this legislative initiative and the policy adaptations that were to follow, the state's Youth Authority quickly established a reputation for its progressive and innovative treatment programs that would soon mark it as a model to be admired and imitated by many other states. In the 1950s, 1960s, and early 1970s, the California Youth Authority pioneered the development of several experimental programs to test ideas that were central to the rehabilitative ideal. For example, the CYA established innovative community treatment projects, diversion programs, probation subsidy services, and youth services bureaus designed to prevent crime and delinquency. Such programs often received widespread attention and high praise from delinquency experts and court officials from other jurisdictions around the county, and such programs were frequently copied by other states.

The CYA maintained its reputation as a progressive and innovative treatment system throughout most of the 1970's. By the early 1980's, however, the CYA's client population and its decision-making processes appear to have changed significantly so

that it would begin to garner a reputation as the "placement of last resort," for the worst of the youthful offenders in the nations most populated state (Little Hoover Commission, 1994).

During the 1980s, the Youth Authority would become a depository, largely populated by what is considered to be "the most serious 5% of the state's youthful offender population" (Skonovd and Haapamen, 1998). The CYA would now house a larger portion of the older, more criminally sophisticated offenders relative to the population that had inhabited its facilities during the first three decades of its operation. Rather than a "model" to be admired around the country, the CYA, in the post-1980 era, would find itself under severe criticism because of overcrowding and budgetary cutbacks. Critics would claim that such conditions severely hinder the CYA's mission to train, educate and treat wards and all but eliminate its involvement in crime prevention activities (see, for example, Lerner, 1982, 1986, 1991; Little Hoover Commission, 1994). Below we provide a brief chronicle of what we regard to have been the most significant legislative and policy changes that were to transform the CYA from its idealistic inception in 1941, through the serious dilemmas it faced in our sample period (1981 and 1992), and continues to face today.

A Short History of Major CYA Policy and Legislative Changes from 1941-1992

From its beginning in 1941 until the late 1970's, the basic criteria for admission to the CYA was whether or not the juvenile or young adult was deemed to be one who could "materially benefit" from the education, treatment, and training services provided by one of its facilities. Length of commitment to CYA institutions was to be limited only

by the age of the offender (by law jurisdiction ended at age 21 for juvenile court referrals, 23 for adult misdemeanants, and 25 for adult felons).

All individuals who were admitted for treatment to the CYA were to have their cases reviewed by the Youthful Offender Parole Board (YOPB) that in 1941, consisted of three members appointed by the Governor. In 1966, membership in the YOPB was expanded to eight persons. Until 1980, the YOPB was part of the Department of the Youth Authority, but legislation was passed that year creating a separate YOPB, one that was independent of the CYA. This legislation also reduced the size of the Board to seven members, with one of its members assuming the role as chairman (State of California, 1993: 2). For the period under study here (1981-92), the YOPB was charged with the formal responsibility of making administrative decisions for all inmates committed to CYA facilities. Among the decisions for which the YOPB assumed overall responsibility were those that dealt with determining the length of each ward's sentence, the return to court of commitment for re-disposition, the specifications of the conditions of parole, the recommendations for the types of treatment programs to be administered to CYA wards, the determination of time until the ward's next Board appearance, and the return of non-resident persons to the jurisdictions of their state of legal residence.

From the 1940's until the 1970's then, young adults and juveniles were all eligible for treatment in the CYA. Beginning in the 1970's, however, case law and legislation were to lead to a number of changes in the eligibility requirements for CYA admission. First, by statute, status offenders were prohibited from CYA commitment, having been declared in 1974 to" no longer materially benefit from the care and custody of secure treatment in CYA facilities" (Cal. Welf. & Inst. Code: 731). Second, and also by statute

(1974), juveniles and young adults charged with the same offenses as incarcerated adults could no longer be confined for longer periods of time than that established by statutory limits for their adult counterparts who were serving time in state prisons (Cal. Welf. & Inst. Code: 1766). In addition, statutory criteria for remanding 16 and 17 year-olds to court for trial as adults were to be relaxed in specific serious cases, thus placing the burden on the juvenile offender to demonstrate that he or she deserved to be treated as a juvenile, and not as an adult offender (Cal. Welf. & Inst. Code: 707). Any juvenile waived to the adult criminal court for trial, however, must have first undergone a CYA Amenability Hearing before they could be bound over to the adult criminal courts for trial (Cal. Welf. & Inst. Code: 707.2). As a result of case law, in 1974, the criminal courts were to be bound by the CYA's determination of amenability, unless these courts could produce substantial countervailing reasons to overturn the CYA's amenability decision on such matters. At the same time, the maximum age of jurisdiction for persons committed to the CYA from juvenile courts for specific serious crimes that were perpetrated at age 16 or 17 was extended by statute from age 21 to age 23. To be more specific regarding this change, there were now three categories of wards that were incarcerated in CYA facilities. First, there were those juveniles who were referred to the CYA from the juvenile courts. These individuals had to be released from the CYA by age 21 or 25, depending on the type of crime(s) for which they were adjudicated. Second, there were those offenders who had been committed to the CYA by the adult criminal courts. These persons had to be released from the CYA by age 25 (or age 23 if they had been referred for a misdemeanor), although sentence enhancements were possible for disciplinary infractions that occurred at CYA institutions, or for parole violations. Third, there were

those individuals who had been sentenced in adult criminal courts to state prison terms, but who were court-ordered to be housed at the CYA. These individuals were to be kept at the CYA until they turned age 25, at which time they were either to be released from custody or transferred to state prisons.

Legislative changes continued in the 1980's. In 1980, for instance, the maximum age of jurisdiction over those referred to the CYA from the juvenile courts for specific serious crimes was raised from the age of 23 to 25, and the proviso that such offenses had to be committed at age 16 or 17 was dropped from the statute (Cal. Welf. & Inst. Code: 1732.6). By the end of 1981, however, many of the legislative changes that dealt with CYA wards began to reflect a dissatisfaction with the rehabilitative ideal and there was a noticeable shift in policy which, for the most part, appeared to emphasize a "get tougher on crime" orientation. For example, at the end of 1981, statutes were amended so that anyone sentenced to a "full" life term for murder at age 18 or older who was court-ordered to be housed at the CYA was no longer eligible for CYA treatment. Policy change toward greater severity of punishment was further evident in 1982, when state statutes were amended so that the courts were no longer bound by the CYA's Amenability Hearing (Cal. Welf. & Inst. Code: 1731.5). Thus, by 1982, the trial courts were no longer bound by the CYA's determination of eligibility for CYA treatment. Therefore, any juvenile who was now waived to, and convicted in an adult criminal court could be sentenced to serve time in a state prison. The CYA's determination of eligibility was now just one factor to be weighed by the trial courts when considering where to remand the convicted offender to custody.

Also in 1982, the voters of California passed Proposition 8. This proposition prohibited the commitment of adults 18 or older (who were convicted of major felonies) from CYA commitment. Prior to the passage of this proposition, approximately 50% of the commitments to the CYA were adult referrals. This enactment immediately reduced the percentage of adult referrals appreciably. In 1983, however (in partial response to Proposition 8), legislation was enacted that permitted some young adults (under the age of 21 at the time of their offenses) who were convicted in adult courts and remanded to state prisons, to be housed in the CYA under court order (Cal. Welf. & Inst. Code: 1766). These individuals were referred to by CYA personnel as the "M cases" or "housing cases."

The economic recession that began in the early 1990's greatly impacted the financial resources of the state of California and, in turn, its ability and/or willingness to provide treatment programs for those incarcerated for their criminal activities. The state's juvenile and adult incarceration rates continued to climb precipitously throughout the 1980's and 1990's, fueled largely by the increase in the severity of punishment meted out by courts to offenders convicted of violent offenses. In fiscal year 1991-92, the CYA was forced to trim nearly $60 million in funding from its budget despite annual increases in the number of wards committed to its care. By June 30, 1992, almost 60% of the CYA's institutional population had a violent offense as their primary commitment offense. Not only were treatment programs cut back at this time, but financial constraints imposed by budget reductions had severely impacted staffing and other program needs at CYA facilities, such as funds for building maintenance and improvements. Again, these budget reductions occurred at a time when there was a substantial increase in the composition of

violent offenders in the CYA's inmate population, and while there was a significant increase in the incarceration of serious drug offenders. As the average length of stay for all offenses among those committed to the CYA increased, and as the resources available to fund treatment programs for these individuals diminished, security risks at CYA facilities accelerated for both staff and inmates alike (State of California, 1993: 7). Violent wards were soon matriculated from one CYA institution to another in an effort to minimize their disruptive influences within these facilities. As a result of these events, judges became reluctant to place offenders who had committed less serious crimes in CYA facilities whenever local alternative treatment was available. Therefore, the youthful offenders who were perhaps the most likely to have benefited from existing treatment programs in the first three decades of the CYA's operation were now being diverted to alternative placement whenever possible. Thus, by the middle of 1992, the CYA would find itself under severe attack by its critics, who accused it of warehousing intractable wards (see DeMuro et al. 1988). Impacted severely by financial limits and cutbacks imposed by the state, and saddled with a population that was dominated by older, more criminally sophisticated violent youth, the CYA was now viewed by its critics as a "post graduate school requiring an undergraduate degree in unsuccessful disposition for admittance" (State of California, 1993: 8).

To document the fact that both the average daily population and the average length of sentence served in CYA facilities increased significantly over the period under study here, we present Table 4.1. The data presented in Table 4.1 show the average daily population and some of the characteristics of all first admissions to the CYA from 1980-1992. This table indicates that there was a general linear upward trend over this period in

Table 4.1. Characteristics of the California Youth Authority and First Commitments, by Year

Characteristic	Year												
	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
CYA													
Total Admissions	3968	4083	3316	2891	3216	3756	3835	3643	3776	3631	3615	3474	3837
Avg. Daily Population	5179	5699	5810	5869	6081	6638	7680	8448	8812	8394	8096	8098	8310
% Capacity	92	102	102	103	103	114	131	146	154	141	132	129	126
First Commitments													
Length of Stay (in months)	12.9	13.1	14.2	15	16.1	17.1	17.8	18.8	21.9	21.6	22.1	21.8	24.9
Court													
Juvenile %	55.2	53.1	67.3	77.2	66.2	58.9	60.7	66.9	68.9	66.7	67.3	70.9	69.4
Criminal (Adult) %	44.8	46.9	32.7	22.8	33.8	41.1	39.3	33.1	31.1	33.3	32.7	29.1	30.6
Commitment Offense:													
Violent %	49.9	49.1	45.0	42.1	41.2	39.8	38.3	35.2	37.4	41.0	47.6	51.3	57.2
Property %	43.5	44.9	47.5	47.7	46.5	47.3	44.2	42.7	38.2	34.9	32.5	31.8	28.4
Drug %	2.3	2.1	2.7	5.6	5.6	7.0	12.0	15.9	18.4	18.6	15.0	11.0	9.1
Other %	4.3	3.9	4.8	4.6	6.7	5.9	5.5	6.2	6.0	5.5	4.9	5.9	5.3
Ethnicity:													
White %	33.7	31.9	32.5	31.1	31.1	34.4	30.5	27.7	25.8	22.9	20.4	18.4	17.2
Hispanic %	28.6	27.5	28.2	30.8	29.8	29.9	30.7	32.4	30.5	33.6	38.6	40.9	46.2
African-American %	35.4	38.0	36.7	35.2	36.0	32.6	34.2	36.0	39.0	37.5	34.3	31.7	29.0
Other %	2.3	2.6	2.6	2.9	3.1	3.1	4.6	3.9	4.7	6.0	6.7	9.0	7.6

Note: Adapted from "California Youth Authority Committee Report: A Study of the Departmental Ward Intake Policy" (State of California 1993), and also incorporating information provided by the Ward Information and Parole Research Bureau of the CYA.
With the exception of the Length of Stay calculation (which is based on first commitments *released* during that year), the characteristics listed in this table refer to first commitments *admitted* to the CYA during the specified year.

both the average daily population at CYA facilities, and in the average length of time served by wards. Comparing 1981 to 1992, for instance, we find that the average daily CYA population went from 5699 to 8310 persons (a 46% increase), while the average length of stay at a CYA facility went from 13.1 months to 24.9 months (a 90% increase). Over this entire period (1981-92), the average daily population at CYA facilities was 124% above its design capacity.

THE CYA: 1981-1992

Between 1981 and 1992, the CYA operated eleven institutions, four rural conservation camps and eighteen parole offices throughout the state of California. At the end of 1992, the CYA had over 5000 employees who were responsible for administering to approximately 15,000 wards and parolees under its supervision throughout the state. Under California law, before a juvenile could be committed to a CYA facility, a judge must first find that there were no other local alternative services or facilities that fit the youth's needs, and that it was probable the he/she would benefit from treatment at a CYA facility (Cal. Welf. & Inst. Code: 707.2). The county probation departments in which these youths lived were to examine the local available alternatives, the youth's prior record, the current commitment offense and other relevant factors, and then make a placement recommendation to the judge. If the judge believed that no suitable local alternative care existed and then recommended a CYA placement for the individual, this decision could be challenged by the youth's attorney, and was subject to appellate review. The youth's probation report and the judge's order were then sent to the CYA intake unit, and based on this material, it was this unit that would decide whether the

youth would likely "materially benefit" from the CYA's treatment facilities. While the CYA was empowered to refuse admittance to any youth it believed might not benefit from its care and treatment, it seldom rejected a committed youth. Youth Authority personnel estimated that less than a half dozen youths per year were turned away by the intake unit (Little Hoover Commission, 1994: 110). In general, the preferred minimum age for admittance to the CYA was listed as 11, but acceptance of wards under the age of 13 was rare and required the approval of the Director of the CYA.

The CYA would, however, sometimes accept youths from counties that had few available treatment facilities, even though these youths were relatively less involved in serious crime than were their counterparts from counties that had available alternative treatment resources. This was especially true in the early 1990s when fiscally strapped counties had to make substantial budget cuts due to the state's financial crises. Many of these budget cuts, however, were to became permanent, and after the financial emergency ended some counties redistributed their limited funds in ways that supported priorities other than local juvenile delinquency, diversion, treatment and prevention programs (Little Hoover Commission, 1994: 58). The California Legislative Analyst's Office, for example, estimated that about 25% of the wards accepted into the CYA programs during the period under study here were "less-than-serious" offenders. These individuals had been referred to the CYA by the twenty California counties that spent little money on local treatment options (Little Hoover Commission, 1994: 110). This indicated to CYA critics that geography rather than individual crime history or individual needs sometimes played a role in determining who was sent to CYA facilities. CYA officials responded to this criticism with the claim that actually less than 20% of first admissions came to the

CYA without a prior adjudicated offense, and that most of the commitment offenses of these individuals were indeed for serious crimes (see Alarcon, 1994: 9).

Recall that once committed to a CYA facility, the length of time served by each youth is to be determined by the Youthful Offender Parole Board. The YOPB employs several criteria to make this determination, including the maximum adult sentence for the same crime for which the youth was charged, the chronological age at which the juvenile justice system loses it jurisdiction over the referred youth (21, 23 or 25 depending on the aforementioned circumstances), and a sentencing guideline adopted by the Board.[1] The sentencing guideline utilizes a grid that classifies wards in one of seven categories based on the seriousness of the crime(s) for which the youth had been committed to the CYA. Each of the seven categories has a different recommendation for the time to be served at the CYA before the youth is eligible to be considered for parole. While the crimes that constitute the seven offense categories and the length of time to be served until initial parole hearing for each category were adjusted at the end of both 1982 and 1986, the scheme in use from 1987-1992 will serve to introduce the reader to the basic framework of the guidelines.

In the 1987 sentencing guidelines in use during the last six years of our study period, Category I included what the YOPB considered to be the most serious offenses. It was comprised of the offenses of those youths adjudicated for murder or a kidnapping

[1] Upon release, parolees from the CYA may have some amount of available confinement time (ACT) left on their sentence. Each parolee's ACT is limited by either their age of jurisdiction (i.e., 21, 23, or 25) or the maximum amount of time an adult convicted for the same offense would serve in the adult criminal justice system, which ever occurs first. Thus, while on parole, the YOPB often times cannot detain a ward for a technical parole violation if they have already served the maximum amount of time an adult would serve for the same offense; this is true even if the ward has not reached the age at which CYA jurisdiction expires

involving substantial injury. Category 2 (the next most severe offense category) included the crimes of voluntary manslaughter, forcible rape, child molestation and kidnap for ransom. Category 3 consisted of the offenses of robbery, mayhem, and burglary with great bodily injury. Category 4 included the commitment offenses of involuntary manslaughter, robbery, burglary with enhancement and narcotics trafficking offenses. Category 5 consisted of the crimes of assault, battery, robbery, and first-degree burglary. Category 6 was comprised of firearms offenses, bomb making, arson, and second-degree burglary. Finally, Category 7, the least serious offense classification, included the violations of auto theft, receiving stolen property, drug possession and all other lesser crimes and parole violations for which youths were referred to the CYA.

The data displayed in Table 4.2 reflect the sentencing guidelines in use prior to and following the two administrative changes that altered requirements for length of time to be served by wards before parole eligibility. That is, Table 4.2 shows the average number of months that wards were recommended to serve prior to an initial parole hearing for the seven offense categories in use between 1980 and 1992. The three separate time periods depicted in this table represent changes in the guidelines in use prior to and after November of 1982 and November of 1986 when the Board implemented adjustments that generally lengthened the amount of time to be served before initial parole consideration for several of the offense categories. The post-1980 administrative changes were fully implemented by the calendar years 1983 and 1987 respectively. The data in Table 4.2 indicate that there was a general increase in time to be served for the most serious crimes (Category 1) over time (i.e. an additional year of time served was added in 1983 and again in1987). However, recommendations of time

Table 4.2. Parole Consideration Date Guidelines (Months Until Parole), by YOPB Category

Years	YOPB Category						
	1	2	3	4	5	6	7
1980-1982	60	48	36	24	15	12	0-12
1983-1986	72	48	36	24	15	12	0-12
1987-1992	84	48	36	24	18	12	0-12

Note. Information was obtained from the summary reports entitled "Youthful Offender Parole Board Initial Appearance Hearings" that are produced each year by the Ward Information and Parole Research Bureau of the CYA.

served before parole eligibility for Categories 2-7 remained relatively stable over this time period. The only apparent change in the recommended time to be served until parole eligibility in categories 2-7 that is reflected in Table 4.2 is the change in the post-1987 period for Category 5, where three months time was added.

While YOPB sentence guidelines were used to establish the "theoretical" baseline for the amount of time recommended to be served before eligibility for parole, the Board is permitted to deviate from the grid recommendation by adding or subtracting sentence time after considering the youth's prior record and any other legally mitigating or aggravating circumstances surrounding the commitment offense (Little Hoover Commission, 1994: 106). The sentence guidelines are thus only to be one consideration when determining the actual length of time the wards are told that they can expect to spend at the CYA. We will return to this point later.

According to CYA critics Paul DeMuro and his colleagues (1988: 5-7), the guideline data depicted in Table 4.2 are highly deceptive and don't tell an accurate story of the average time actually served by wards before they are considered for parole. DeMuro and his associates contend that due to the aforementioned sentencing revisions passed at the end of 1982 and 1986, parole consideration dates actually increased for about 75% of the specific offenses listed in the CYA's baseline guidelines. These critics claim that in affect, after the revisions were enacted, the average length of stay doubled for approximately 38% of the listed guideline offenses and stayed the same for only 26% of these offense types. For example, DeMuro and his colleagues (1988: 6) calculated that as a result of the 1987 revisions:

> although the baseline was increased in only two categories

> (1 and 5), about half the offenses were moved into a higher (more serious) category. This resulted in increased time from commitment to parole consideration and greater Board supervision of many offenders. For example, a youth committed for first-degree burglary prior to 1987 would have been a Category 6 offender, serving one year to initial parole consideration. Under the new guidelines, this same youth has become a Category 5 (more serious) offender, and cannot be considered for parole until 1 1/2 years had been served (about 28% of Youth Authority juvenile court commitments are for burglary). Similarly, an armed robber was moved from Category 4 to3, with a 25% increase in time to parole consideration date (these youth account for about 16% of the juvenile court's first commitments).

The YOPB's (1988: 60) response to criticisms of this policy offered by DeMuro and his colleagues was a written reply contending that the decision to revise sentence lengths for certain crimes upward was consistent with the wishes of the legislators, law enforcement personnel, and citizens of the state of California for whom the Board was appointed to serve. In particular, the YOPB maintained in this reply that when considering guideline deviations of a ward's parole consideration date, they consistently followed criteria prescribed by Title 15, Division 4.5, Section 4945 (j) of the California Code of Regulations that had been ratified by the state legislature. These regulations specify that the length of stay at the CYA is to be determined by one or more of the following twenty individual considerations:

(1) Protection of the public.
(2) Prior probation or parole failure.
(3) Attitude and sense of responsibility toward commitment offense.
(4) Attainment of institutional goals.
(5) Institutional behavior.
(6) Participation in program.
(7) Educational potential
(8) Employment potential

(9) Emotional adjustment
(10) Staff evaluation, treatment team report, psychiatric report.
(11) Special psychiatric/psychological needs.
(12) Alcohol/ drug dependency.
(13) Family support.
(14) Future plans.
(15) Placement potential.
(16) Community reaction.
(17) Availability of community-based program to further treatment and training needs.
(18) Motivation and prognosis for success.
(19) Probability of recidivism.
(20) Continuing (or abstaining) participation in youth gang activities while incarcerated.

Given the criteria in Title 15, the YOPB (1988) argued that its change in add-on-time to parole release date was entirely consistent with the state's regulatory guidelines.

DeMuro and his colleagues (1988: 63-64) responded to this reply by the YOPB, however, by noting that the guidelines had led to extraordinarily long sentence lengths that were counterproductive for some offenders. Table 4.3 shows the average sentence length that wards were told they would have to serve before they were eligible for an initial parole hearing for the years 1980-1992 after time was added to or subtracted from the guideline recommendations due to one or more of the aggravating or mitigating legal considerations listed above. That is, this table depicts the average time that offenders in each of the offense categories were told by the Board that they would have to serve before they would be eligible for their initial parole hearing. In lieu of the comments by DeMuro and his colleagues, this table is more instructive than is Table 4.2 when depicting the year-to- year variations in the sentences handed out by the YOPB for various offense categories over time. Table 4.3 indicates the presence of substantial increases for the most serious crimes (Categories 1 and 2) after the administrative

Table 4.3. **Average Months Until Parole at Initial Parole Consideration Date (PCD) Hearing with the YOPB, by YOPB Category and Year of PCD Hearing**

Year of Hearing	All Categories	All First Commitments	YOPB Category 1	2	3	4	5	6	7
1980	16.0	NA	56.5	44.5	31.3	22.2	14.9	13.8	10.4
1981	16.9	NA	59.2	46.7	34.1	23.1	15.4	13.8	11.1
1982	16.0	NA	59.0	47.8	34.5	23.5	15.9	13.7	10.9
1983	16.0	18.4	68.8	52.7	35.0	24.0	16.3	14.5	12.8
1984	15.7	18.6	67.5	48.6	33.0	23.7	16.9	14.4	12.6
1985	16.4	19.4	70.0	42.2	34.1	24.6	18.3	15.3	13.4
1986	16.0	19.6	68.2	48.5	36.0	23.7	18.6	16.6	14.2
1987	23.3	24.0	81.5	50.0	37.5	27.2	21.2	16.9	12.8
1988	19.4	22.8	82.4	48.1	36.9	24.8	18.7	15.1	12.4
1989	18.5	22.1	81.8	46.6	34.9	22.9	16.7	12.8	9.8
1990	19.4	24.5	84.1	47.4	36.0	23.5	17.8	14.0	11.9
1991	19.6	23.9	83.1	45.1	34.9	22.8	16.9	12.9	11.2
1992	19.3	22.8	79.6	44.6	34.6	22.3	16.0	12.1	10.7

Note: Information was obtained from the summary reports entitled "Youthful Offender Parole Board Initial Appearance Hearings" that are produced each year by the Ward Information and Parole Research Bureau of the CYA.

changes were implemented in 1983 and again in 1987. There appears to be a general pattern in Table 4.3 where the wards in these two categories were told to expect to serve substantially longer periods of time than were the wards sentenced under previous guidelines after these two administrative changes were in place. Overall, however, this general pattern of increase in expected time served seems to have been directed only for Category 1 offenders. There was considerably less variation over time in the time that wards in categories 2-7 were told they could expect to serve between 1980-1992.

We believe that the amount of time that wards were told they would have to serve prior to a parole hearing is an important factor because it has a direct effect on their individual behavior while they are in the institution. Of course, we are ultimately less interested in the length of time that wards were told that they could expect to serve before they received an opportunity to gain their release through parole, and most interested in the amount of time that they actually did serve before they were released. Table 4.4 shows the actual average length of time (in months) served prior to release on parole (1980-92) for seven different specific offense types. The reader can see that in many cases wards were paroled earlier than they were told to expect to be released. The general trend in this table, however, indicates an overall tendency to keep wards incarcerated for longer periods over time, especially for the more seriously regarded offenses.

A significant claim made by Demuro and his colleagues contends that the physical structure of a substantial portion of the CYA facilities did not provide adequate protection for those wards that were confined there. For example, these investigators estimated that about 40% of CYA facilities were aging, overcrowded dormitory style buildings. They argued pointedly that because there was a large proportion of

Table 4.4. Average Length of Stay (in months), by Commitment Offense Type and Year of Release

Year of Release	First Degree Murder	Aggravated Assault	Enhanced Robbery	First degree Burglary	Auto Theft	Narcotic Sales	Narcotic Possession
1980	32.1	NA	18.7	12.6	11.0	14.6	12.7
1981	38.1	17.2	19.0	12.9	10.8	10.2	10.8
1982	41.1	19.3	19.9	12.8	11.8	15.1	11.4
1983	53.5	20.3	21.4	15.2	12.4	14.9	12.4
1984	49.5	21.3	24.6	13.1	13.3	17.1	11.6
1985	51.7	23.7	25.0	14.7	13.7	20.9	13.7
1986	55.2	25.6	28.2	17.1	15.7	14.2	15.8
1987	62.8	27.2	28.3	19.9	18.1	18.3	19.0
1988	63.8	30.7	36.9	23.7	20.3	23.7	21.0
1989	72.0	32.1	35.3	23.3	17.3	25.8	21.7
1990	74.5	29.0	34.3	23.2	16.5	25.5	23.4
1991	76.9	28.8	32.7	22.6	16.6	25.1	22.6
1992	74.5	28.2	31.5	21.8	14.5	25.9	21.7

Note Information was obtained from the summary reports entitled "Length of Stay of Youth Authority Wards" that are produced each year by the Ward Information and Parole Research Bureau of the CYA. First Degree Murder in 1980 was for all types of murder

incarcerated violent offenders at the CYA (many of whom belonged to rival gangs) and who were permitted to interact freely within a dormitory setting, that the potential for individual violence was great. This placed both wards and staff at high risk.

Table 4.5 displays the reported assault rate against staff by wards, the assault rate against wards by other wards, the total assault rate by wards, and the total disciplinary infraction rate against wards-- each per 100 average daily population at CYA facilities from 1981 through 1992. With the possible exception of the assault rate by wards against staff, there does not appear to be an upward general linear trend in reported violence at CYA facilities from 1981 through 1992. The data in Table 4.5 also indicate that there does not appear to have been the proportional increase in violence that critics predicted based on the additional crowding over time. These figures, however, do appear to back up the previously mentioned allegation that there was a great deal of recorded violent assaults at CYA facilities during this period. The rate of violence that took place in CYA facilities during this period is high by any correctional standard.

In sum, detractors of CYA policies have maintained that the high rates of assaultive violence and the high disciplinary infraction rates displayed in Table 4.5 reflect largely the freedom of movement and interaction that were permitted by the dormitory style living environment at some CYA facilities. The critics argued that this style of living quarters served to encourage such behaviors. As a result, these critics contended that the dormitory type of living situation was sometimes indirectly responsible for the increasing length of sentences served by some CYA wards (Little Hoover Commission, 1994: 116). More specifically, such detractors claimed that dormitory style environments made assaults and other rule infractions against others more likely. Thus both the high

Table 4.5. Assault Rate Against Staff, Assault Rate Against Wards, Total Assault Rate, and Total Disciplinary Infraction Rate, by Year

	Rate Per 100 Average Daily Population			
Year	Assault Rate Against Staff	Assault Rate Against Wards	Total Assault Rate	Total Disciplinary Infraction Rate
1980	1.6	14.9	16.5	100.6
1981	1.6	9.4	11.0	101.8
1982	1.6	11.8	13.4	116.8
1983	1.9	15.6	17.5	134.9
1984	1.6	15.4	17.0	133.1
1985	1.5	20.1	21.6	142.5
1986	2.2	21.8	24.0	141.5
1987	1.4	12.3	13.7	140.7
1988	2.7	14.6	17.3	122.1
1989	3.4	14.8	18.2	136.6
1990	3.1	14.2	17.3	139.5
1991	3.3	15.0	18.3	143.3
1992	3.8	15.0	18.7	164.6

Note: Information provided by the Ward Information and Parole Research Bureau of the CYA. Assaults refer to assaults by other CYA wards.

population density and the physical structure of the dormitory style buildings was said to increase the likelihood that occupants would assault one another, or would act out by violating CYA rules. These infractions, in turn, were said to often result in added time to ward sentences during annual reviews of their cases by the YOPB. As we discuss below, wards were also given additional sentence time by the Board when they appeared not to cooperate with the demands of treatment programs prescribed for them, or had otherwise failed to complete Board-ordered programs prior to their release date. This policy would also lead to severe criticisms from the critics of the CYA.

In general, critics have maintained that the CYA's ability to provide treatment, education, and training for wards has continually eroded since the early 1980s (see, for example, DeMuro et al., 1988). This erosion appears to have been the result of several factors. First, there had been a steady change in the composition of the ward population as a greater proportion of inmates in the post 1980 era was more frequently and seriously involved in crime than was the case in earlier periods. Second, there had been a general shift in the orientation of both the legislature and the public away from the rehabilitative ideal toward a desire for more severe punishment for law violators, and the decisions of the YOPB reflected this change. Third, budget cuts in the early 1990s had also impacted the CYA's ability to provide beneficial treatment to the ever-increasing number of wards admitted to its facilities. One trend that was especially disturbing to the CYA's detractors is particularly worthy of mention below.

As we indicated earlier, it was the seven member YOPB that set the initial parole consideration date for individual wards shortly after their arrival at CYA facilities. The Board also ordered a specific treatment program for each newly arrived person. The CYA

recommendations may have included counseling, substance abuse programs, educational programs, victim awareness education, parenting skill classes and the like. Each year the Board would evaluate the ward's progress. CYA policy detractors maintain, however, that problems frequently occur when wards are not able to gain admittance to treatment programs that they have been ordered by the YOPB to complete due to the limited availability of space in such programs at CYA facilities. There are often long waiting lists for many of the programs offered at these facilities. When CYA wards cannot enter and complete required programs within the original commitment period due to the limited availability of space, critics contend that their sentence time is routinely increased by the YOPB during the annual reviews of their cases. Hence, critics have charged that confinement times of wards are often lengthened for reasons other than their misbehavior or refusal to cooperate by actively participating in prescribed treatment programs at the CYA (see DeMuro et al, 1988). Critics have maintained then that sentence lengths are often extended because there was simply not enough space available in the very treatment programs that the Board had ordered the youths to complete as a condition of their release from the facility. In fact, in their evaluation of CYA practices, the Little Hoover Commission Report (1994: 106- 109) stated that this "is the single most important factor behind institutional overcrowding" in CYA facilities."

On average, juvenile CYA residents spent approximately 4.3 months longer in confinement than did adults housed in CYA facilities (the M cases) for similar offenses (DeMuro et al., 1988: 9). With respect to sentence length then, adult inmates had an advantage over juveniles housed in CYA facilities because they automatically had one day removed from their sentence for every day they served without being written up by

staff for commission of a disciplinary violation. Because juveniles housed in CYA facilities did not automatically receive these "good time" cuts. DeMuro and his colleagues (1988: 8) report that wards adjudicated in juvenile courts resented this inconsistency, and saw it as just one more example of the injustices inherent in the CYA system.

The CYA's detractors also claim that the YOPB has too much control over ward treatment decisions. The critics contend that as the YOPB's role in making treatment decisions expanded in the 1980s, the role that the professional staff played in such decision-making has diminished proportionately. Critics maintain that the professional staff at the CYA has thus become increasingly estranged from the YOPB. During the period under investigation here many of the CYA staff were said to have believed that the Board had too frequently acted independently of their judgments, and did not pay enough attention to their recommendations. Due to their frequent contact with wards and their clinical experience, many CYA staff were said to believe that it is they who were most qualified to make these program decisions. Increasingly then, the critics maintained that such decisions were made independently of CYA staff input by Board members who were often without benefit of extensive clinical training, and who saw these youths only for about 10 to 15 minutes each year. As a result, the critics argued that many CYA counselors were left to feel like they "now run a high-security warehouse for people for which they have little to say over who comes, who leaves, or what they do while they are there." (DeMuro et al, 1988: 7).

The YOPB responded in writing to these charges (1988: 60) with the assertion that "such criticisms were very much exaggerated." In their response to their critics, the

YOPB estimated that CYA staff followed the placement and program recommendations of the Board only about 60% of the time, and that in actuality, the majority of the Board's placement decisions were indeed based on staff recommendations and input. Thus, the YOPB contended that "the majority of the differences in placement decisions was due to lack of bed space in the agreed upon placement rather than any philosophical disputes of treatment and training programs." The critics" addressed this response by the YOPB by claiming that the Board was "buck passing" because it admits that there are not enough treatment options at CYA facilities (see DeMuro et al, 1988:64).

Rehabilitative and Parole Services Available at CYA Facilities (1981-92)[2]

According to literature provided by the Youth Authority, during the time period between 1981 and 1992 several specialized and supplementary programs were offered at its facilities in its efforts to transform wards into productive law-abiding citizens. For example, over this entire time period the CYA attempted to enhance the potential job-related skills of all its wards. It mandated therefore that every ward participate in a course that averaged between 6 to 8 weeks that was designed to develop the individual's employability skills. This course emphasized how to develop self-awareness, employment goals, on-the-job skills, and career awareness. It also aimed to teach the ward how to fill out job applications and design a resume, and how to find jobs and prepare for job interviews.

The more general core programs available at each of the CYA facilities during this period consisted of three main components: education, treatment and training

programs. Below we describe briefly each of these core components of the CYA's program.

Education

The vast majority of CYA wards are school dropouts or underachievers. The CYA thus places a great emphasis on improving educational skills. Individual wards were ordered to improve their educational skills to a certain specified level as a pre-condition to be eligible for parole release. All of the CYA's education programs were competency-based, which means that wards must have achieved specific definable and measurable outcomes before they could advance through the curriculum. The program was structured so that wards may have been enrolled in more than one program at the same time. The major components of the education program were as follows:

a. Middle School-- this component offered instruction in basic education and in the development of citizenship skills to wards who were 13 years of age or younger.

b. Basic Skill Enhancement-- this component offered remedial language, math and reading instructions to those wards 14 years of age or older who were considered to be underachieving. Those enrolled in this program received elective high school unit credit for work completed and were allowed to take concurrently any high school courses for which they had met the prerequisite requirements.

c. Career Vocational Preparation—this component provided students at all levels of core programming with pre-vocational and vocational training in a number of careers choices including Food Services, Auto and Body Repair and Maintenance, Computer Repair, Welding, Cabinet Making and Mill Work.

d. High School—this component provided curriculum standards and required courses endorsed by the State Board of Education. Completion of this program permitted the

[2] The material for available rehabilitative and parole services between 1981 and 1992 was adapted from Written Testimony for the Little Hoover Commission. Department of the Youth Authority (March, 1994).

CYA to recommend the granting of a diploma that had credibility with educational institutions, the community, and employers.

e. College- this component permitted qualified students to pursue a college education by enrolling in college correspondent courses.

In addition to the basic core educational program outlined above, each CYA facility also provided wards with special supplementary services that were designed to enhance competency in specialized areas of the core program by making available additional resources to augment ward proficiency with basic educational skills. Supplementary Services available in each of the CYA facilities included:

1. Special Education--this component assisted wards with specific learning disabilities, severe emotional disorders, physical, speech and language impairments or handicaps.

2. English as a Second Language—this components offered special classes or special attention to wards in regular classes whose primary language was not English.

3. Elementary Secondary Education Amendment (Chapter I)--this component provided additional assistance to those wards that were below par in their reading, language, and math skills.

4. Adult Basic Education—this component provided remedial programs for wards over the age of 16 who were not enrolled in programs that earned high school credits.

5. General Educational Development Test—this component provided testing for wards in order to obtain a certificate that was equivalent to a high school diploma (GED).

6. Educational Counseling—this component offered students guidance, behavioral counseling and career employment counseling.

7. Job Training Partnership Act--this component offered government funding which augmented part of the vocational training at several of the facilities. Supplementary support for this program was provided by the Private Industry Council to improve job-training skills for wards.

8. Vocational Education Act—this component supplied additional grant monies authorized specifically for criminal offenders that would expand, improve, and produce innovations in vocational programs at CYA facilities.

9. Library Services—this component supports library services at each institution that would provide for the acquisition of recreational reading, audio and visual resources, and other materials related to development of curriculum and treatment programs and/or the expansion of the legal library at facilities.

Treatment

Each of the wards housed in CYA facilities was required to participate in treatment and counseling programs. These programs were designed to assist wards in understanding the causes and the consequences of their misbehavior and tried to ensure that such behavior would not occur in the future. Because most of the wards had long histories of frequent and serious illegal behavior, the CYA had developed a number of comprehensive specialized programs that were available at some but not all of its facilities. Wards were initially diagnosed at CYA clinics and then (when possible) were sent on to those facilities that had the specialized treatment programs that were appropriately suited to handle his/her particular needs. Upon arrival at the facility, each ward was assigned a Counselor who was to conduct regular individual and small group therapy sessions with the ward. Treatment goals and objectives were established by these counselors for each of the individual wards. The ward's progress was then monitored by a Case Conference Committee of counselors who were responsible for reviewing every aspect and component of the individual's treatment program during his/her stay at the CYA facility.

In addition to individual and small groups counseling, a number special and supplementary programs were also available at specific CYA facilities. Below is a

description of the treatment programs available at certain CYA facilities between 1981 and 1992.

a. Intensive Treatment Programs (ITP)— were designed for wards determined to have serious emotional problems. ITP programs integrated psychotherapy with the other elements of the wards core programming. Such wards resided in a specially trained living unit away from the general population throughout their stay. They were counseled and cared for by a team of psychiatrists, psychologists, nurses, interns and volunteers. ITPs were located at four of the CYA's facilities around the state.

b. Specialized Counseling Programs (SCP)—SCPs were designed for CYA wards who had been adjudicated for crimes such as sex offenses, violent acts, and arson. These programs were tailored to the specific needs of the wards, and were staffed by psychologists and consulting psychiatrists. Wards participated in intensive therapy sessions with a psychologist and in group therapy sessions with youth counselors. These programs were designed so that they could be completed within one year. Those who considered to have successfully completed the program were then integrated into the CYA's general population.

c. Substance Abuse Programs. —each of the CYA facilities had a substance abuse program designed to intervene in the ward's use of alcohol or drugs. All wards believed to have substance abuse problems must have participated in this program. These programs ran from six to twelve months, and wards generally could not be considered for parole release until they completed this program.

d. Planned Re-entry Program (PREP)—located at the Ventura facility, the PREP program was a short-term intensive-counseling program for those relatively unsophisticated offenders who did not have long prior records of involvement in serious crime. The program generally ran from five to seven months, depending upon how long it took for the ward to achieve specific individual treatment goals.

Training

All CYA activities and programs were structured toward training wards to acquire skills and values with the aim of helping them to become productive and law-abiding citizens. CYA programs emphasized values such as self-discipline, positive thinking, high self-esteem, a strong work ethic, and personal responsibility. Wards were held responsible and consistently disciplined for improper and illegal behavior at institutions.

Attempts were made to teach self-control through crisis intervention techniques, and wards were expected to show good personal hygiene and to keep their living units clean. They were also to be provided with positive role models through associations with staff and volunteers from the community. Below are some of the specific training programs available to wards during the period under study.

a. Free Venture Programs—since 1985, the CYA has co-sponsored a program with private industry and public agencies that set up a business or services within state facilities, and than trained and employed wards to manufacture products or perform services. Private and public companies equipped and operated the business and provided supervision and mentoring to the wards. Wards were hired by these agencies through an interview process and they were expected do demonstrate values, skills, attitudes and behavior that met the standards set by these companies. Example of such programs are the TWA reservations service at the Ventura school, and the fluorescent ballast quality control industry located at the DeWitt Nelson facility in Stockton. Wards typically received wages equivalent to an apprentice salary or starting-level salary. During the period covered in our study, twenty percent of the ward's net earned wages went to the state to compensate for the cost of incarceration. Another 15% of the wards gross pay went to the state victim compensation program or to the victim of the particular ward's crime if a restitution order was in effect. And forty percent of the ward's net salary went to a savings account that was available to wards upon their release. The remainder of the ward's paycheck went to the individual's trust fund account at the facility and was used to make approved purchases within the CYA facility.

b. Public Service—during the period covered on our study all CYA facilities had a public service component that attempted to instill a sense of responsibility in wards and emphasize the obligation to donate time to community service. The CYA estimated, for example, that wards provided 545,400 hours of public services to California communities in 1992. The CYA facilities that contributed most heavily to the public service program were the camps that had long been engaged in wild land restoration, park cleanup, fire fighting, and related forms of land conservation. Other CYA facilities performed services that ranged from road maintenance and highway cleanup, to public land restoration and the repairing Christmas toys. The Karl Holton School in Stockton had perhaps the most acclaimed public service unit in the state. Its Mountain Public Service unit had received recognition from the Governors office and the National Association of Search and Rescue. Twenty wards in this unit provided 24-hour emergency services to several counties that included search and rescue for lost, stranded and injured citizens. These wards also regularly taught survival skills to

elementary school children and renovated and removed litter from campsites and woodland areas.

c. Citizen Participation Programs--all CYA facilities depended on voluntary citizen participation programs to augment their curriculum. Each separate facility had its own Volunteer Service Program Manager whose job it was to recruit, train, schedule and supervise community volunteers to mentor, tutor, visit, provide recreation, conduct religious services, counsel, and offer other supportive services to wards that were not possible through normal budgetary resources. The CYA believes that such services helped to bond institutionalized wards with the outside community.

d. Impact of Crime on Victims Program--implemented in 1984 at all institutions and camps, the CYA conducted regular classes in the impact of crimes on victims and the consequences of driving while intoxicated. The goal was to increase the ward's sensitivity to the victim's plight and to teach them the relationship between early participation in crime and later criminal behavior. Victims of crime were frequently invited to discuss the consequences that crime has had on their lives.

While the menu of programs described above suggests that there were a number of diverse educational, treatment and training programs available to wards at various CYA facilities during the period covered in our study, several CYA employees who read this chapter stated that the discussion of services offered implies more programming than was actually available at these facilities. For example, several employees contend that it is ludicrous to suggest that all wards had counselors who conducted regular individual and small group therapy sessions. Critics have also maintained that in practice, overcrowding prevented many youths from enrolling in available programs within the bounds of their initial prescribed sentence length. This overcrowding was, according to such critics, primarily the result of YOPB policies (DeMuro et al. 1988). The critics maintained that when wards were finally able to gain entrance into these programs there was little chance that the curriculum would succeed in rehabilitating them. This was said to result because daily life at CYA facilities had become a contest for self-protection and survival where

any potential gains from individual or group programs were quickly cancelled out by the conditions imposed by the violent sub-culture that permeated these institutions. The YOPB responded to such criticisms with the admonition that any sentence length time additions incurred by wards had more to do with the ward's behavior and gang associations at the CYA than with YOPB policy changes (YOPB, 1988: 60).

Parole Service Programs

In addition to education, treatment and training programs, the CYA also had an extensive parole service program to assist wards when they were released from CYA facilities. When the ward was deemed eligible for parole by the YOPB, a report would be forwarded to a parole casework supervisor who would then review the file and make a determination as to the level of supervision and the type of services that would be needed to supervise the youth in the community. The goal was to promote the successful integration of the individual into the community while at the same time protecting the public from a possible return to crime by the ward. A pre-release conference would be established between the ward and the staff in person or by telephone in order to review relevant case information and to plan a program that would meet the ward's individual needs. At this conference, the terms and conditions of the ward's parole and the level of the required supervision and services would be determined. The CYA's parole release system was based on a "step down" model in which the level of supervision and services would be reduced over time as the ward successfully met the conditions of his/her parole. The major options that were available for parole supervision between 1981-92 are listed below:

a. Electronically Enhanced Parole Release Program (EEPRP)--this program was initiated to reduce institutional crowding by releasing selected wards (those who were not adjudicated for serious violent offenses) 60 days earlier than their original release date into a highly structured parole supervision program. Wards released into this program were required to wear electronic monitoring devices and were not permitted to leave their residence except for pre-authorized departures to seek employment, attend school, or receive treatment or counseling. Parolees were randomly contacted by supervisors at least once a week and were drug tested a minimum of twice per month. The "in house" arrest system saved the state money and the contacts with parole agents and electronic monitors provided for 24 hours surveillance. When violations were suspected, response was immediate. Minor violations were sanctioned by loss of privileges and temporary detention. Serious infractions were treated as parole violations and generally led to revocation of parole. Those wards that successfully completed the 60-day period were next assigned to either the intensive re-entry or specialized caseload programs described below, depending on their perceived needs.

b. Intensive Re-Entry--eighty percent of the parolees released by the CYA received this service. Those who did not receive this service lived in geographic areas where it was impossible to provide this supervision. Here a parole agent tried to coordinate the actions of the parolee, family members, employers, teachers and relevant community organizations in order to ensure as smooth a re-entry as was possible. This program generally ran from 75 to 90 days for most cases, and over 90 days for the most serious offenders. In this program the parolee received a minimum of two contacts per week for the first month of release from his/her parole agent, and weekly contacts for the duration of the re-entry period. Those who had a history of drug and/or alcohol problems were tested twice monthly for indications of substance abuse. During the re-entry period, the parolee received employment, education, or job training assistance, individual and group counseling, and subsidized placement services as needed and/or available.

c. Specialized Caseloads--in this program parole agents were assigned reduced caseloads relative to those of the regular (case management) parole officers. These special parole agents were expected to give intensive, concentrated attention to parolees who had particular needs such as persons with mental problems, sex offenders, street gang members and substance abusers. This program attempted to both increase the likelihood of a successful parole adjustment for such individuals and to make the earliest possible determination of potentially dangerous behavior on the part of the parolee. Parolees remained on intensive re-entry status until such time as they exhibited a stable pattern of behavior for a significant period of time in order to demonstrate to the parole supervisors that they were no longer in need of intensive care, and were not a danger to public safety. If these conditions were satisfied, the parolees were then transferred to a case management parole program.

d. Case Management--those who successfully met the conditions of intensive-re-entry and specialized caseload parole programs were transferred to a case management parole officer. These parole officers assisted the parolee in maintaining acceptable levels of behavior and job stability. Case management parole agents had higher caseloads than did the specialized agents described above, and were less concerned with providing services to parolees and more concerned with monitoring their illegal behavior. They did so by making unannounced substance abuse testing visits and unannounced visits to the homes and work places of parolees. Prior to their release, parolees were individually classified by case management agents as "maximum," "medium," or "minimum" risks. If classified as a maximum risk, the parolee would be contacted a minimum of twice per month by a case management agent. Medium risk parolees were contacted once per month, and minimum risk parolees received one contact every other month. Case reviews were mandated at specific intervals for all parolees and wards so that they could be reclassified as higher or lower risks based on any new information obtained by agents.

California law obligated that the "Parole staff shall assist parolees in obtaining adequate housing, employment, financial assistance, social and medical services, educational placement, and other resources or services which will increase the likelihood of a parolee's adjustment in the community." Twenty percent of the Parole Branch's operating budget had been allocated to subsistence and personal care services for parolees during the period under study. In many cases, assistance for living arrangements upon release on parole must have been initiated before the ward left the institution. Many parolees were without personal resources or had no family or friends that would assist or agree to assist them upon their release. Parole agents assisted in finding living placement alternatives for these individuals, usually among motels, foster or groups homes, YMCA rooms, or if available, residential treatment centers that had facilities for 24-hour supervision and intensive services.

The CYA also operated some smaller parole service programs. For example, two residential parole programs for technical parole violators with substance abuse problems

were developed to deal with the problem of institutional overcrowding. Both programs offered the option of volunteering to undergo 90-days residential care treatment in lieu of serving between six months to a year after being returned to the CYA. The CYA also operated two intensive supervision service programs located in residential facilities in San Diego. The CYA named these facilities the NETWORK program because it was a highly specialized contracted residential placement that provided a wide range of interconnected services that included psychiatric treatment and employment services, as well as educational, vocational and recreational treatment programs.

As was the case for education, training and treatment programs, critics would also maintain that the potential success of the parole services program was adversely affected by changes in the policies and practices of the YOPB that led to stricter standards for parole violation in the 1980s. It was the contention of these CYA detractors that the raising of these standards led to higher parole revocation rates. For example, in 1986 the Board mandated that all technical violations of parole must thereafter be reported directly to it. Critics would argue that once the YOPB was supplied with this information, it was too quick to revoke parole, even for non-criminal technical violations. DeMuro and his colleagues (1988: 7), for example, reported that in 1986, 25% of all CYA admissions were for technical parole violations, and that in 1987, this number increased to 37%. Critics argued, therefore, that the YOPB had become increasingly strict in its standards such that the focus of parole had changed drastically over time period studied here. DeMuro and his colleagues (1988: 8), report that parole officers in the past had concentrated on providing parole services for those released from the CYA in order to

keep them from recidivating. Since 1986, however, "the parole officer's operating policy had changed to what one high-ranking CYA officer had termed trail'em, nail'em, jail'em."

Responding to such criticisms, the YOPB' (1988: 60) replied that much of the increase in parole revocations could be accounted for by the commission of new felony offenses by parolees. They contended as well that most of the Board's actions were actually consistent with the wishes of parole officers who themselves were increasingly recommending revocation of parole for a cumulative series of incidents and/or lesser offenses committed by parolees. The Board added that any indications of drug and gang behavior were especially likely to lead to parole revocations. The Board's contention then was that it was for reasons that mainly had to do with parolee behavior that parole revocations had increased, and not because of any other specific change in Board policy.

In sum, it is evident from the material presented above that there was a sincere attempt to offer many program options and parole services to CYA wards during the period covered in our study. Recall, however, that critics maintained that the CYA facilities were so overcrowded that wards would often have difficulty obtaining space in these programs and that YOPB policies themselves were a direct cause of this overcrowding. Furthermore, critics also claimed that even when wards were able to find space in the various treatment programs in which they had been ordered to participate by the YOPB, the hostile and dangerous environment in which these programs were delivered would offer only minimal treatment value at a high expense. In the long run then, the critics contend that CYA programs afforded little protection for the public's safety. According to such critics, while the long sentences imposed by the YOPB served to temporarily incapacitate CYA wards, keeping them from committing crimes against

the public, Californians were not getting much of a long-term return on their investment. While incarcerated, many wards were becoming more embittered and more criminally sophisticated and were likely to commit new crimes soon after they were released on parole.

To YOPB members, however, changes in sentence length for serious offenses were in line with what was happening in society. Changes in Board policies that affected the length of time served by wards had in fact kept crime rates low by incapacitating hard-core criminals who had high propensities for serious crime. The Board's policies simply reflected the public's and politician's wishes to be tough on such individuals. In addition, contrary to the claims of the critics that the CYA staff was demoralized because the YOPB had failed to follow their treatment recommendations, Board members maintained that any policy changes they directed toward wards were in fact, consistent with staff recommendations and input.

Having described at length the history, policies and programs of the CYA, we now turn our attention to describing the data we use to conduct subsequent analyses.

CHAPTER 5

DATA AND METHODS

INTRODUCTION

As argued in Chapter 3, one of the major contributions this study can make to the literature on continuity and discontinuity in criminal offending patterns is derived from the nature of the data to be used in the forthcoming analyses. No other published research to date analyzes such a large sample of this nation's most serious youthful offenders over such a long period of time. This chapter details the methods and data to be used in the analyses presented in Chapters 6, 7, and 8. Again, our data consist of information on three samples of males released from the California Youth Authority in fiscal years 1981-82, 1986-87, and 1991-92.

The first section of this chapter describes the data sources, the specific variables to be used in the datasets, and the limitations of these data. This section also contains a description of how the final analytic samples were constructed. We conclude this chapter with a discussion of the statistical methods employed in this study; this section will include a description and comparison of the finite mixture or semiparametric random effects models of Nagin and Land (1993; Land and Nagin 1996; Land et al. 1996) and parametric random effects models.

Before we describe the data and methods to be used here, let us briefly review and embellish where needed, the process of becoming a ward of the CYA. Recall that during the period of time in which the three samples employed in this study where under supervision, youthful offenders were committed to the CYA as one of three broad

commitment types (State of California 1993): juvenile court, adult court, and the adult state prison system. *Juvenile Court Commitments* were juveniles between ages 11 and 17 at the time the criminal offense occurred, adjudicated in juvenile court, and then committed directly to the CYA from the juvenile court (State of California 1993). *Adult Court Commitments* occurred in one of two ways, either as juveniles remanded to adult court or as young adults committed directly to the CYA. Juveniles originally remanded to and convicted in adult criminal court could have been found amenable for treatment in the CYA if they were not sentenced for an offense carrying a life term.[1] Young adults between the ages of 18 and 21 could have been found amenable for treatment in the CYA. Both of these commitment types from adult courts were still considered "regular YA" cases, were assigned a normal 5-digit CYA number and were subject to the jurisdiction of YOPB. In 1982, however, voters of California passed Proposition 8 that prevented direct commitments from the adult criminal court for those offenders between the ages of 18 and 21 who were convicted of a "serious felony." Before the passage of this proposition, about 50% of the commitments to the CYA were adult referrals. The passage of this proposition served to reduce the percentage of the adult court referrals.[2]

---

[1] Whether a juvenile will be remanded/waived to the adult system depends on the outcome of the "fitness hearing," as prescribed by Section 707 of the California Welfare and Institutions Code. It is this hearing that determines whether the juvenile will be treated as either an adult or a juvenile. In 1982, the public became concerned that the transfer of juveniles to adult court was too complicated, and thus a series of legislative changes in 1982 changed the nature of the fitness hearings. The legislative changes to Section 707 of the California Welfare and Institutions Code meant that juveniles who were over the age of 16 and accused of serious/violent felonies were now deemed to be *unfit* for treatment in the juvenile justice system if they failed on any of the following five factors (Humes 1996): (1) criminal sophistication; (2) potential for rehabilitation; (3) previous delinquent history; (4) success of previous attempts at rehabilitation; and (5) the gravity of the current offense(s).

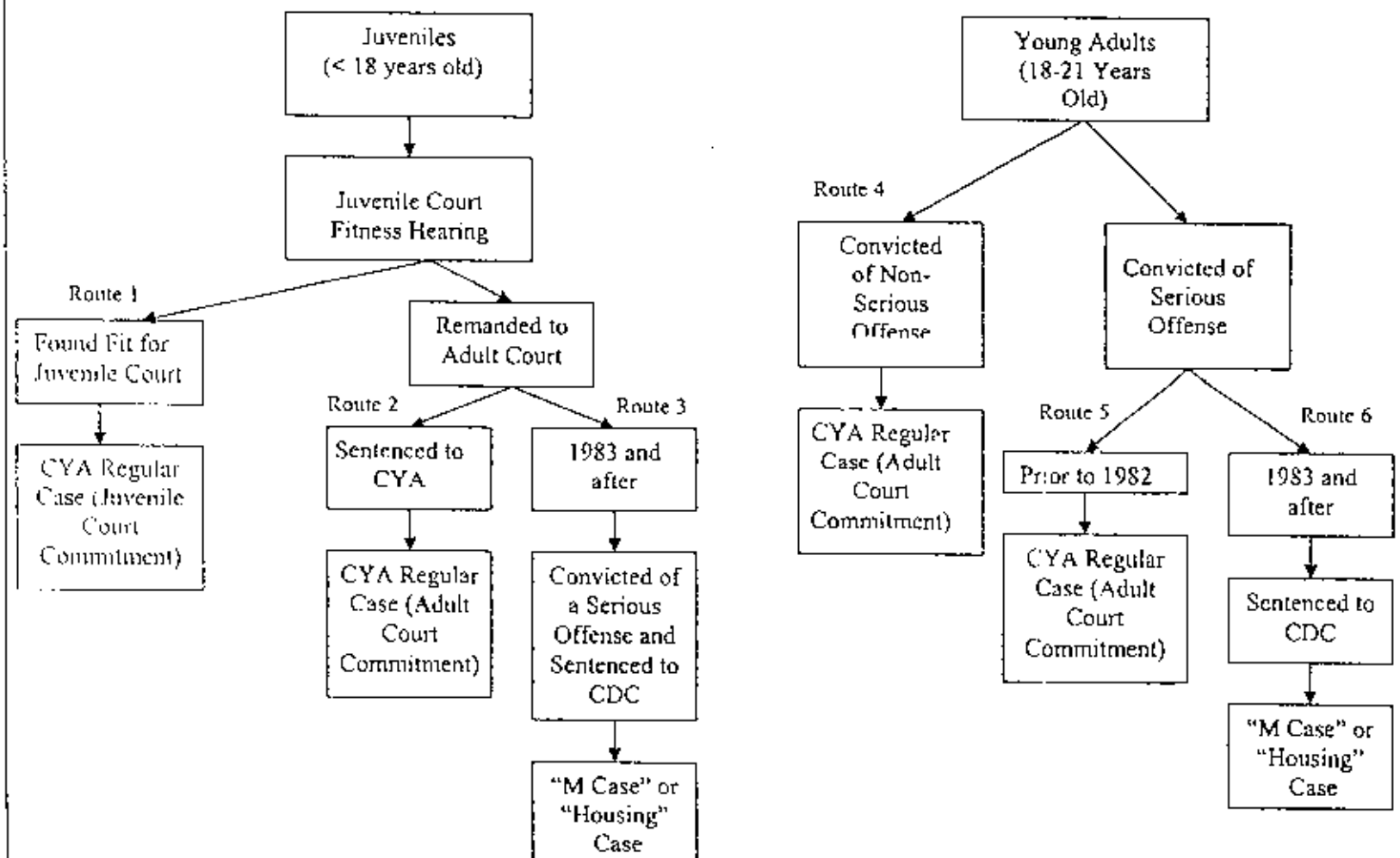
[2] For example, in 1980 the CYA population was composed of 55.2% juvenile court commitments and 44.8% were adult court commitment. By 1985, the corresponding percentages were 58.9 and 44.2, and in 1995 they were 81 and 19 (adult court commitments here include both straight commitments and M Cases) (State of California 1993; Legislative Analysts Office 1996). By December 31, 2000, the population was

Finally, *state prison commitments housed in the CYA* (known as *M Cases*) refer to a category of offender who was under the age of 21 (including both adults and remanded juveniles) at the time he/she was *sentenced to* the California Department of Corrections (CDC). These offenders from adult criminal court were "housed" in the CYA if they met a number of conditions, including (but not limited to) space availability, availability of adequate facilities, no history of aggressive or assaultive behaviors in prior correctional programs, not already a parolee of the CYA, not having previously been found non-amenable to treatment by the CYA, and not previously discharged as a ward of the CYA. At any point in time, an M Case could be immediately transferred to the California Department of Corrections (CDC) if they were found to be either a threat to institutional security or "intractable." The YOPB did not have formal jurisdiction over the M cases (or housing cases) and thus these offenders did not make formal appearances before the YOPB. This possible pathway to the CYA began in 1983 as a response to Proposition 8, that explicitly prohibited the commitment to the CYA for offenders sentenced in adult court for the commission of a serious felony (e.g., felony index crimes).

Figure 5.1 depicts six possible routes into the CYA for both juvenile and young adult offenders in our samples. The far left path in Figure 5.1 (Route 1) depicts the pathway for the juvenile offenders found "fit" for juvenile court and committed directly to the CYA. The other two pathways reflect how juveniles (under the age of 18) found to be "unfit" for regular juvenile court and "waived/remanded" to adult court ended up in the CYA. Prior to 1982, juveniles remanded to adult court could have subsequently been

95.4% juvenile court commitments, with 3.4% straight adult court commitments and 1.4% M Cases (Legislative Analysts Office 2001).

Figure 5.1 Graphical Depiction of Routes of Entrance into the California Youth Authority



committed to the CYA as a direct adult court commitment regardless of the commitment offense (excluding any offenses carrying a life sentence), as long as they were perceived to materially benefit from treatment in the CYA. This pathway is the middle pathway for juveniles depicted in Figure 5.1 (Route 2). The right-most pathway for juveniles depicts one that opened up in 1983 as a response to Proposition 8 that explicitly prohibited the commitment to the CYA for offenders sentenced in adult court for the commission of a serious felony offense (e.g., Index crimes). Thus, one segment of the "M case" population during this time period were juvenile offenders (less than age 18 at the time of a serious offense) who were found "unfit" for regular juvenile court. Such individuals were then remanded to and convicted in adult court and subsequently sentenced to the CDC and then finally ordered housed in the CYA (Route 3).

The right-hand side of Figure 5.1 depicts three routes to the CYA for young adult offenders. First, young adult offenders convicted of a relatively non-serious offense in adult court could have been directly committed to the CYA as an adult court commitment (Route 4). Otherwise, if one was sentenced before or after Proposition 8, young adult offenders could have been committed directly to the CYA if they were deemed amenable to treatment (Route 5) or they could have been "housed" in the CYA if they met strict regulations and qualified as an "M" Case (Route 6). Incidentally, legislation passed in 1996 all but eliminated this final path of M Case classification for young adult offenders.[3]

In addition to being classified according to whether a ward is committed from the juvenile court, adult court, or as an M Case, wards in custody of the CYA are also

[3] This legislation limits M Cases housed in the CYA to those under age 18 at the time of sentencing or those that have a potential release date set by the CDC to be prior to age 21. Prior to this legislation, M Cases could have been housed at the CYA until the age of 25. M Cases that will not be released before age 21 are now automatically transferred to the CDC at age 18 (Legislative Analysts Office 2001).

classified on the basis of three broad additional categories: (1) *first commitments*, (2) *parole violators*, and (3) *recommitments*. First commitments are wards committed to the CYA for the first time. Parole violators are wards previously incarcerated in the CYA, then released on parole, and who have had their parole revoked for a violation(s) of the terms of parole (e.g., arrest for a criminal offense, gang activity, positive drug test, failure to be employed, AWOL). Finally, recommitments are wards previously committed to the CYA, released from the CYA, and then subsequently recommitted to the CYA again for a new criminal offense.

Having reviewed background information on the CYA as an institution and identified the various pathways by which an individual can be incarcerated in the CYA, we can now turn to describing the three release samples that constitute the data sets employed in the analyses of this study.

THE THREE RELEASE SAMPLES

The data used in this study consist of three samples of California Youth Authority wards. Norman Skonovd and Rudy Haapanen of CYA's Research Bureau had previously collected the data for the two earliest samples (1981-82, 1986-87) (Skonovd and Haapanen 2000); the data for the 1991-92 sample were collected by Michael Ezell, Lawrence Cohen, Norman Skonovd and Rudy Haapanen (with funding for the data collection provided by the National Institute of Justice).[4] To maintain consistency across the three samples, the 1991-92 sample was coded according to the same rules and procedures used in collecting the two prior samples. The only differences in the initial

[4] NIJ Grant Number #98--CE--VX--0026.

coding rules and procedures employed were with the data on prior criminal history (as described below).

The two earliest samples are based on data collected on 2,086 wards released from the California Department of the Youth Authority institutions in fiscal year 1981-82 (July 1, 1981 – June 30, 1982) and 2,078 wards released in fiscal year 1986-87 (July 1, 1986 – June 30, 1987).[5] Initially a random sample of 2,200 wards was drawn from the 4,425 wards released from Youth Authority institutions in fiscal year 1981-82 and another 2,200 wards from the 3,048 wards released in fiscal year 1986-87. Of the 2,200 wards in the 1981-82 sample, 114 cases were removed from the study because their records were "court-ordered seals" which prevents any access to their files.[6] This resulted in a 2,086ward data file for the base-line data set. Of the 2,200 wards in the 1986-87 sample, 122 cases were removed for the same reasons. This resulted in a 2,078 ward data file for the second data set. Together, these two samples provide records on 4,164 individual wards.

The arrest data for these two samples originally were only available through December 31, 1991 for the 1981-82 sample and December 31, 1990 for the 1986-87 sample. We have, however, updated the arrest data for these two samples through June 30, 2000. This adds about ten years of additional arrest data to these two samples.

The 1991-92 sample is based on data collected on 1,527 wards released from the California Department of the Youth Authority institutions in fiscal year 1991-92 (July 1,

---

[5] Throughout this study, the term "*sample stay*" is used to refer to the incarceration period in the CYA that resulted in the ward being released during the sampling time frame. The date of release for the sample stay was the key defining element that resulted in the ward being included in one of the samples

[6] When a CYA ward's records are ordered "sealed" by the court, the ward's CYA Master Files are sealed and the ward's 5-digit "YA number" is replaced with an S and 4 digits (e.g., S0001). Law then prohibits access to the ward's juvenile offending history and CYA Master File.

1991 – June 30, 1992). Initially, a random sample of 2,198 wards was drawn from the total of 4,030 wards released from Youth Authority institutions in that fiscal year.[7] Of the 2,198 cases, 13 of the sample members were subsequently removed because their records were "court-ordered seals" (resulting in an initial file of 2,185 cases).

Importantly, prior to beginning the formal coding on the 1991-92 sample, concerns arose over possible time and budget limitations given that cases were taking longer to code at the culmination of the training sessions than previously estimated. The decision was made that guaranteeing accuracy of the coded cases would be of greater value compared to the speedy collection of data for the entire 2,198 cases. Accordingly, a random number was assigned to each case in the sample at the outset of the formal coding process, and cases were then coded according to their random number (rank ordered from lowest to highest). This ensured that at the end of the available time allotted and available financial resources, the resulting sample would still be a random subset of the initial sample. At the point in time when both the time and financial resources had been exhausted, 1,527 of the original 2,198 (70%) cases had been coded. Thus, the final file for the 1991-92 sample consisted of a total of 1,527 wards. Comparisons of the wards who were coded with the wards who were not coded indicated no significant differences in terms of ethnicity, county of commitment, commitment offense, court of commitment, and the probability of arrest after release from the CYA.[8]

---

[7] The sample initially had 2200 cases in it, but two individuals were sampled twice because they had been released twice during the sampling time frame. We only allowed each of these individual's one record in the data file, and we selected the later release date for these two individuals. The later release data was selected so as to not artificially create a very rapid failure time, which would have been assured if I had chosen the earlier release date.

[8] The comparisons were made based on the fact that some variables were available regardless of whether the case was coded because certain information is always gathered and maintained by the CYA on all

The fact that the coded cases did not appear to be significantly different from the remainder of the initial sample on any critical variables is not surprising given that 70% of the initial sample was coded and that the cases were strictly coded according to their random numbers; thus, the "randomness" of the file was maintained even though the entire sample initially drawn could not be coded.

We have also updated mortality data for the two earliest samples and collected original mortality data on the 1991-92 sample though December 31, 1999. The possible mortality of the subjects in this study is important for two reasons. First, we did not want a ward to appear as "arrest free" simply as a function of their death. Thus, wards who had died were removed from the risk of arrest after the point in time at which they were found to have been deceased. Second, as argued by Gottfredson and Hirschi (1990: 94), individuals with low levels of self-control should experience "death at higher rates than the general population." Analyses by Dobrin (2001) showed (not surprisingly) that individuals with criminal arrests in their backgrounds have a greater chance of dying by homicide compared to individuals with no criminal arrests in their backgrounds. Preliminary analyses of the previously collected mortality data on the two earliest data sets analyzed herein lend some credibility to this assertion as well (see Lattimore et al. 1997).

---

wards who are committed to their facilities. This information was available electronically through the OBITS (Offender Based Institutional Tracking System) computer system of the CYA. Further, since the arrest data for the period of time after the sample release was obtained electronically, we did obtain the post-release criminal history data for all of the original 2,186 cases (that had available post-release criminal history data)

DATA SOURCES AND VARIABLES

The sources of information that were used to collect the data varied according to both the type of data element and whether the data element was referring to a characteristic or behavior *prior to* or *after* the date of release from the CYA that resulted in the ward being included in the sample. We will refer to these two different segments of the data collection process as the "pre-release" and the "post-release" periods. The key division point is the date of release for the "sample stay," which refers to the specific incarceration or "stay" at the CYA that resulted in the ward's inclusion in one of the three samples. The "pre-release" period refers to time prior to the sample parole release, while the "post-release" period refers to time after the date of release.

Pre-Release Data: Case Characteristics Information

Data for the pre-release period on the characteristics of the cases were collected from two sources: (1) Youth Authority's electronically stored information on the ward and (2) Youth Authority's "hard copy" ward Master Files. From various computer files within the CYA and the CYA's Offender Based Institutional Tracking System (OBITS), data were obtained for the following variables:

- *Date of admission and release for sample stay*
- *Base commitment offense* (e.g., adjudication for murder, forcible rape, burglary, robbery, grand theft auto)
- *Admission status* (first commitment, parole violator, recommitment)
- *Date of birth*
- *Gender*

- *Ethnicity* (White, African-American, Hispanic, Asian, Native American, Filipino, Pacific Islander, Other)
- *Court of commitment* (juvenile court or adult criminal court)
- *M Case* (CDC "Housing" Case)
- *County of commitment* (e.g., Los Angeles, Sacramento, San Diego, San Francisco, Alameda)
- *Major CYA infractions* (known as DDMS violations) for such things as fighting, rioting, assaulting another ward or staff member, gang activity, and drug use.

The second data source is the individual, hard copy Master File completed for each ward. The Master File contains all available prescribed program and parole data, as well as data pertaining to the ward's entire medical, educational, psycho-social, and criminal history up through the date of discharge from Youth Authority's jurisdiction. Records concerning the ward's behavior and characteristics are required to be included with the ward's commitment papers by California Welfare and Institutions Code #1741:

> the judge before whom the person was tried and committed, the district attorney or other official who conducted the prosecution, and the probation officer of the county, shall obtain and with the order of commitment furnish to the authority, in writing, all information that can be given in regard to the career, habits, degree of education, age, nationality, parentage and previous occupations of such person, together with a written statement to the best of their knowledge as to whether such a person was industrious, and of good character, the nature of his associates, and his disposition.

Experienced coders reviewed and coded relevant data from the following types of documents: police, probation, and court reports, Youth Authority staff reports and documents, consultant reports and evaluations, and letters and appeals. Information from these sources was coded according to uniform guidelines. The Master File was the major source of information regarding the prior behavior and characteristics of the cases.[9] Information regarding prior criminal record, as well as family background, substance abuse, gang activity, and prior placement information is either not available in CBITS or it is not as complete or accurate as that contained in the Master File. The following variables were coded from the detailed information contained in the Master Files:[10]

- *Family violence*: Evidence of violence among the family members (not including the ward).
- *Parental alcohol or drug dependence*: Evidence that the ward's parents or guardians have an alcohol or drug dependence problem (e.g., arrests for drug offenses, been in treatment for drug/alcohol problems). Social drinking or occasional marijuana use was not recorded as evidence.
- *Parental criminality*: Evidence the ward's parents have been previously involved in criminal activity (e.g., prior arrests or incarcerations). One arrest for drunk driving was not recorded as evidence.

---

[9] The Master Files that were coded ranged from several hundred pages (smaller Master File) to several thousand pages (larger Master Files). One Master File in the 1991-1992 sample contained 11 inches of paper reports.

[10] In the following descriptions, "parental" indicates the ward's parents unless the parents are no longer the guardians of the ward. In that case, parental refers to the ward's guardians.

- *Sibling criminality/delinquency*: Evidence the ward's siblings have been previously involved in criminal or delinquent activity (e.g., prior arrests or incarcerations). One arrest for drunk driving was not recorded as evidence.
- *Parental lack of supervision/neglect*: Evidence that the ward was not adequately supervised or ward was neglected by his or her parents (e.g., ward was removed from custody of parents due to the behavior of the parents; "parents do not know where the ward is usually").
- *Ineffective parental control*: Evidence that the parent's had ineffective or inconsistent control over the ward (e.g., ward arrested for being "out of control" or noted in probation as being "out of control"). Naturally, since all of the wards in these samples were arrested for criminal offenses at least once, arrests for criminal offenses were not recorded as evidence of being beyond the control of the parents.
- *Physical abuse*: Evidence that the ward has been subject to either extreme punitiveness or physical abuse (e.g., parent arrested for abuse of the ward; severe whippings; spankings that cause injury). Spanking alone (without injury) was not recorded as evidence.
- *Sexual abuse*: Evidence that the ward has been subject to sexual abuse by others (e.g., molestation, intercourse with adult persons, adjudicated sexual abuse case).
- *Drug abuse*: Evidence that the ward abuses drugs (not including alcohol). Experimental drug or chemical use was not recorded as evidence of drug abuse. Daily or frequent use of hard drugs such as cocaine, PCP, and heroin, and "sniffing" were recorded as evidence.

- *Gang member/association*: Evidence that ward associates with gang members, participates in gang activities for self-protection, or is a fully participating "gang member" that engages in "gang banging." Often could be identified by the presence or nature of their "moniker" (e.g., "Little OG"), previous arrests for gang-related activities such as drive-by shootings, or the presence of gang tattoos that denote affiliation (many of the tattoos were on the neck and hands—"18th Street" was tattooed across the forehead of one ward).
- *Previous violent behavior*: Evidence that ward has previously been violent, including assaultive behavior and arrests for violent offenses.
- *School dropout*: Evidence the ward has dropped out of school. Evidence for this variable included the ward not being enrolled in school, ward had not attended school for six months (even if they have not formally "dropped out") while free on the street, ward had been expelled, or the ward was persistently truant (e.g., absent without excuse more than they are present).

Information on four other variables was also collected, but these variables were highly collinear with the other variables described above. Since these variables merely included redundant information, they were not considered further in this study. These variables included alcohol abuse, school disciplinary problems, violent behavior while in the CYA, and gang activity in the CYA. The specific variables that were highly correlated were (1) drug abuse and alcohol abuse, (2) school dropout and school disciplinary problems, (3) previous violent behavior and violent while in CYA, and (4) gang membership and gang activity in CYA.

Pre-Release Data: Arrest History Information

The arrest history of each ward prior to the sample stay was also compiled using the information contained in the Master Files. As noted above, the Master File contains all of the previous probation reports and court records of wards because these data are legally required to be submitted with the court order of commitment to the CYA. Using all of the available police reports and records, probation reports, court records, and CYA parole performance summary information (for the wards previously released from the CYA prior to the sample stay), the following variables were coded and checked for accuracy for each arrest event:

- *Date of arrest event*: The date the ward was arrested by law enforcement personnel. In the rare situation (<1% of the time) when that arrest date was not known, the date of the offense was coded.
- *Arrest Charges*: Up to 3 arrest charges per arrest event were coded (i.e., some arrest events involved multiple charges against the offender). Only behaviors that reflect distinct law violations were coded as separate charges and only the most serious charge per behavior was coded; "lesser-included" offenses were never coded. For example, if a ward was arrested for evading the police in a stolen car after a robbery, the three arrest charges would reflect each behavior (robbery, auto theft, evading the police). In the cases where a ward is arrested for grand theft auto, occasionally the "lesser-included" offenses of possession of stolen property, and unlawful taking of a motor vehicle were also filed. We only coded the most serious arrest charge per behavior, and thus this arrest would reflect a single arrest

charge (grand theft auto). Similarly, if a ward was arrested for attempted murder, as well as assault with a deadly weapon, and assault and battery, we only coded the attempted murder arrest charge (e.g., you can't attempt to kill someone with a firearm without committing both assault with a deadly weapon and assault and battery). Allowing multiple arrest charges per arrest date is a more accurate way of cataloging an individual's prior record (Geerken 1994).

As noted above, the only difference in coding procedures and rules between the two earlier samples and the most recent sample was in the coding of the criminal history data. In actuality, all distinct arrest charges for a given arrest event were coded for the 1991-92 sample, not just the three most serious charges. This allowed us to ascertain if there were any biases associated with using only the three most serious charges rather than all arrest charges. First, 93% of the arrest events had three or fewer arrest charges; 98% of the arrests only involved four or fewer arrest charges. Second, of the charges that were dropped, over 70% involved only charges for drunk in public, possession of alcohol, giving false information to a police officer, being under the influence of a controlled substance, and other "miscellaneous" relatively minor charges. There was a precipitous drop off in the seriousness of the arrest charges after the third arrest charge, and in no case did dropping these records result in a ward being misclassified as a nonviolent or nonserious offender. Third, among the 1,460 males that were coded in the 1991-92 sample, the mean number of charges was 9.62 using only the three most serious charges, whereas it was only 9 8 if we allowed for all of the arrest charges. Thus, using only the

three most serious charges seems to accurately depict the arrest histories of this sample with little possibility of bias.

To make the prior arrest data of the 1991-92 sample equivalent to the two prior samples, we employed the same process used in coding the earlier two samples to arrive at the three most serious charges per arrest event. Although that data were coded manually and then entered into a computer database, we automated this process using a computer program that looped through the arrest charges for each event and pulled out the three most serious charges. First, all offenses were classified according their corresponding OBITS offense category (which ranges from 1-100), as was performed in the two prior samples. Then, the computer program looped through and pulled out the three most serious charges according to the seriousness hierarchy that was programmed into a computer algorithm. Any charges that were ranked fourth or lower according to the algorithm were then dropped. In all of the analyses in this study, only the three most serious charges per arrest event were employed for the 1991-92 sample.

Briefly, the algorithm always considers violent offenses the most serious charges, then serious property offenses (e.g., burglary, auto theft), followed by major drug offenses (e.g., sales and trafficking), and, finally, the least serious miscellaneous charges (e.g., petty theft, drunk in public, trespassing). Appendix A contains a table listing the seriousness hierarchy of the offenses.

Post-Release Data: Arrest History Information

The source of data for arrests that occur after release from the CYA for the "sample stay" is the Automated Criminal History System maintained by California

Bureau of Criminal Statistics and Criminal Identification of the California Department of Justice (CDOJ). This data source was used to obtain the post-release criminal history data because neither OBITS nor Master Files contain any relevant arrest data subsequent to each ward's respective discharge from Youth Authority (which is usually, but not always, subsequent to a period of parole). The data from this third source are known as the California Information and Identification "CII rap sheet" information. A list of CII identification numbers for the wards was submitted to the CDOJ, who then compiled a data file containing all of the information in the CII rap sheets (including arrest records) of the wards in our samples.

When an individual is committed to the CYA, the ward is assigned a CII identification number and a computerized (automated) CII rap sheet file is initiated and maintained by the CDOJ. When an adult is arrested in California, the arrest is reported by the arresting law enforcement agency to the California Department of Justice (which houses the state repository for arrest data). Thus, any time one of the wards in the samples was arrested as an adult, the arrest record including the date of arrest and information on the arrest charges was forwarded to the state repository of these data. If a ward was released by the CYA while still a minor (under age 18), the CYA reported any subsequent criminal arrests of the ward while he or she was a minor to the CDOJ.

The files of the California Department of Justice were searched in late November of 2000. We permitted five months of "lag time" to allow sufficient time for any arrests to be entered by the Department of Justice into the case's "rap sheet" file. Thus, the arrests were censored as of June 30, 2000 and any arrests occurring between that date and

November of 2000 were not included in the analyses for this study.[11] The post-release exposure periods for the samples were between 18-19 years (depending on the date of release) for the 1981-82 sample, 13-14 years for the 1986-87 sample, and 8-9 years for the 1991-92 sample. The average ages at the end of follow-up period (June 30, 2000) were 37, 33, and 27 for each of the release samples, respectively.

To make the post-release data equivalent to the prior arrest data, we included only the three most serious charges per arrest event. We extracted the three most serious charges using the same process described above for the 1991-92 prior arrest data. We extracted the following variables from the CII rap sheet data files:

- *Date of arrest event*: The date the ward was arrested by law enforcement personnel.
- *Arrest Charges*: The 3 most serious arrest charges per arrest event.

Post-Release Data: Mortality Information

Mortality data on the cases in the release samples were extracted from the Death Statistical Master Files (DSMF) of the California Department of Health and Human Services (DHS). The DSMF files are based on the death certificates completed by either the presiding physician at the time of death, or in the case of sudden or unexpected deaths such as homicide, suicide, or drug overdose, by the coroner or medical examiner investigating the deaths. There is one DSMF file for each year. For example, all of the deaths that occurred between January 1, 1990 and December 31, 1990 would be included in the 1990 DSMF file. We had access to the DSMF files for 1989-1999 and thus the last

[11] There were 402 arrest events that occurred between July 1, 2000 and November 30, 2000.

known possible date of death for our data would be December 31, 1999. Death dates prior to January 1, 1989 for the 1981-82 and 1986-87 samples were obtained from data previously compiled by Skonovd and Haapanen (2000). Mortality data are crucial for the topics of this study since they remove an individual from being "at risk" of arrest when no longer alive. Thus any cases that died after release will not counted as individuals who were "arrest free" at any age purely as a result of their mortality. As will be shown in Chapter 6 of this study, a sizable number of cases died (usually of homicide) after release from the CYA.

From the information contained in the DSMF files, we retrieved the following two variables:

- *Date of death*
- *Cause of death*: International Classification of Death (ICD) codes were used to identify the major cause of death (e.g., homicide, suicide, drug overdose, accident, auto accident, and AIDS).

Appendix B describes the process used to obtain the dates of deaths; this process allowed for the identification of any deaths that occurred in the 1981-82 and 1986-87 samples between 1990 and 1999, and to gather death data through December 31, 1999 for the 1991-92 release sample.

Post-Release Data: Adult Incarceration Information

Due to the fact that the CII "rap sheet" files only contains accurate reports of the dates of *intake* into the state penal system (California Department of Correction—CDC),

with the help of Lee Britton, Norman Skonovd, and Rudy Haapanen of the CYA Research Bureau and Christopher Haws of the CDC, recently we were able to obtain the adult incarceration records related to all of the stays in the CDC subsequent to release from the CYA. Information on the following variables was made available:

- *Date of intake*
- *Date of release*
- *Commitment Offense*: Indicates the criminal offense that resulted in the case being incarcerated in the CDC.
- *Second Strike*: Indicates if the case had been sentenced as a "Second Strike" case.
- *Third Strike*: Indicates if the case had been sentenced as a "Third Strike" case.

Deriving the Analytical Samples

In order to be included in the analyses in this study, there were several conditions a case had to satisfy. Table 5.1 details the effects that adhering to the conditions for inclusion in the final analytical sample had on the final sample size.

The first constraint used in deriving the analytical sample was the gender of the ward; the analytical sample was limited to only male wards. This constraint was imposed for two primary reasons. The major reason the female wards were excluded was that there were simply too few females in each of the datasets to allow for separate models or reliable estimation of model parameters. Table 5.1 indicates that females constituted only

Table 5.1. Constructing the Analytical Samples: Limiting the Release Samples by Sex, M Case Status, and Missing Arrest Data, by Release Sample

Limit Imposed	Release Sample								
	1981-82			1986-87			1991-92		
	Original	Starting	Removed	Original	Starting	Removed	Original	Starting (N=1527)	Removed
Step 1: Sex									
Female	88 (4%)	88 (4%)	88	81 (4%)	81 (4%)	81	87 (4%)	67 (4%)	67
Male	1998 (96%)	1998 (96%)	0	1997 (96%)	1997 (96%)	0	2098 (96%)	1460 (96%)	0
Total	2086	2086	88	2078	2078	81	2185	1527	67
Step 2: M Case Status									
Yes	0 (0%)	0 (0%)	0	208 (10%)	203 (10%)	203	220 (10%)	0 (0%)	0
No	2086 (100%)	1998 (100%)	0	1870 (90%)	1794 (90%)	0	1965 (90%)	1460 (100%)	0
Total	2086	1998	0	2078	1997	203	2185	1460	0
Step 3: Missing Pre Arrest Data									
Yes	0 (0%)	0 (0%)	0	206 (10%)	151 (8%)	151	---	0 (0%)	0
No	2086 (100%)	1998 (100%)	0	1872 (90%)	1643 (92%)	0	---	1460 (100%)	0
Total	2086	1998	0	2078	1794	151	---	1460	0
Step 4: Missing Post Arrest Data									
Yes	8 (0.4%)	7 (0.4%)	7	18 (12%)	200 (12%)	200	71 (3%)	26 (2%)	26
No	2078 (99.6%)	1991 (99.6%)	0	1830 (88%)	1443 (88%)	0	2114 (97%)	1434 (98%)	0
Total	2086	1998	7	2078	1643	200	2185	1460	26
Step 5: Missing Death Date									
Yes	3 (0.001%)	2 (0.001)	2	0 (0%)	0 (0%)	0	0 (0%)	0 (0%)	0
No	2083 (99.9%)	1989 (99.9%)	0	2078 (100%)	1443 (100%)	0	2185 (100%)	1434 (100%)	0
	2086	1991	2	2078	1443	200	2185	1434	26
Final Analytical Sample		1989			1443			1434	

Note: The numbers in bold indicate the frequency (percentage) of cases at that are available at the "start" of the next limiting step.

about 4% of each sample (n=88, 81, and 87 for the 1981-82, 1986-87, and 1991-92 samples, respectively).[12]

A lesser reason why the females were excluded is that there is much empirical evidence that shows that male and female offending patterns are not equivalent (for example, see D'Unger et al. forthcoming). Thus, we preferred to exclude the female cases entirely. It was clear during the coding of the 1991-92 sample that there was a marked division between the offending patterns of the males and females, with the offending patterns of the male wards indicating significantly more frequent, more serious and more violent behavior. This is not to say that the females in the original samples were not serious and/or violent, but just that compared to the males in the sample they were not as violent and/or serious offenders. Comparing all of the males in the samples to all of the females, the mean number of prior criminal arrest charges was 9.80 for the males, and it was 6.8 for the females. Similarly, 90% of the males in these three samples were arrested in the post-period, whereas 76% of the females were arrested. However, the offending patterns of the females in these samples compared to typical females in the general population are certainly both much more frequent, more serious, and more violent; this was especially true for the female wards who were gang members. Nonetheless, the female cases were removed at this point, and the sample sizes at the end of this step were 1,998 (1981-82), 1,997 (1986-87), and 1,460 (1991-92) respectively.

The second constraint imposed for inclusion was that only the cases that were "regular" CYA cases (i.e., directly committed to the CYA) would be included in the final

[12] Removing the females with missing arrest data further reduced these already small samples. Further, recall that only 1,527 cases were actually coded in the 1991-92 sample. Of these 1,527 cases, only 67 of them (4%) were females.

analytical sample. Thus, M Cases were not included in the final analytical samples. Recall that M Cases are the California Department of Corrections commitments "housed" in Youth Authority facilities (i.e., juveniles who have been "waived" to the adult criminal justice system or select young adult offenders). Due to the fact that M Cases are not subject to the YOPB control, they do not have extensive clinic summaries and Youthful Offender Parole Board-related records in their Master Files. Thus, because the M cases are not Youth Authority commitments they are not subject to supervision by this parole board. Since they were not subject to overview by the YOPB, their Master Files are missing the reports and documentation that contain the necessary information for some/many/all of the background characteristics recorded here. Thus their offending patterns were not described in the detail like the regular CYA cases (that are subject to the YOPB supervision). It was impossible to code these cases with the same detail and attention that was given to the CYA regular cases.[13]

As shown in Table 5.1, roughly 10% of the initial 1986-87 and 1991-92 samples were M Cases released from the CYA institutions during the respective fiscal years. As

[13] It also bears noting that there are not only selection effects that determine what cases would end up in the CYA as M Cases, but since all M Cases are subject to immediate transfer to the CDC if they misbehave and threaten institutional security, it is likely that there are even heavier selection effects determining what M Cases are released from the CYA in a given year. The M Cases that would be released from the CYA would be the best-behaved M Cases—the problem M cases would have already been transferred to the CDC. In fact, drawing on our experiences while we were at the CYA collecting the 1991-92 sample, we would suspect that the M case population at intake is not very different from the general CYA population (which is why the end up being "housed" there in the first place), but the M cases that do not get transferred to the CDC and complete their sentence while housed in the CYA would have marginally better chances to remain arrest free after release than would the regular CYA cases (because of the selection effects determining what M Cases are left at the end). In fact, since the post-release (follow-up) arrest data was gathered electronically, we did have the post-release arrest data for both the CYA regular cases and the M Cases, and analyses of failure rates for the two groups in both samples with M Cases supported this position. For the male regular CYA cases released in 1986-87, 91% of them had been arrested at least once by June 30, 2000, whereas 85% of the male M Cases in that sample had been arrested. For the 1991-92 cases, the corresponding arrest percentages as of June 30, 2000 were 89% (male regular CYA cases) and 80% (male M Cases)

noted above, both budget and time constraints were a problem with the coding of the 1991-92 sample, and as a result the decision was made that the M Cases would only be coded pending available time and budget resources after coding of all regular CYA cases. This decision was made after extensive discussion with CYA Research Bureau personnel with considerable experience with CYA Master Files and after we reviewed the CYA Master Files of 10 randomly selected M Cases. Our review of these 10 cases indicated that they could not be coded with the same accuracy and detail (including the arrest histories) as the regular CYA cases. Thus, none of the 220 M Cases in the original 1991-92 sample was coded. As shown in Table 5.1, after removing the M Cases (that were coded) from the 1986-87 sample that entered this step, the resulting sample size of the 1986-87 sample was now 1,794 (1986-87); the samples for the 1981-82 and 1991-92 samples were unaffected by this step either because there were no M Cases in the CYA population (1981-82) or because the M Cases were not initially coded (1991-92).

The next two constraints required for inclusion in the final analytical sample concerned whether any of the "pre-release" and "post-release" arrest data were missing. Cases were not included if they were missing the prior criminal arrest history. This turned out to be a major problem only for the 1986-87 sample. Of the 1,794 male, regular CYA cases that entered this step for that sample, 151 (8%) were found to be entirely missing their arrest histories. Skonovd and Haapanen (2000) did not find any apparent pattern of bias related to whether or not a case was missing their prior arrest history. We did not code these data, and it is beyond our speculative powers to assess why these cases are missing their arrest data. Our analysis of the probability of arrest in the post release period indicated that the cases missing their prior arrest histories were

more likely to remain arrest free in the post-release period than were those cases with available prior arrest data. Eighty-six percent of the cases missing their prior arrest data were arrested after release, whereas 92% of the cases not missing their data were arrested. Thus, both of these groups experienced what might be called "excessive failure rates," but one group was marginally more excessive than the other.

Cases also were not included in the final analytic samples if we did not have access to their CII "rap sheet" arrest data for the post-release period. There are a variety of reasons why an individual's CII identification number would not be available at a point in time after their release. Sometimes a ward's CII Identification Number (which is attached to their fingerprint) is not entered into the OBITS system because it is not known or available, and occasionally records are purged from OBITS (as a result of discharge from the jurisdiction of the CYA) prior to retrieving their CII number. The OBITS system was used to obtain the CII identification numbers that were submitted to the CDOJ, and thus if the CII identification numbers were missing in the OBITS system, the CII rap sheets for those wards could not be obtained. Missing CII identification numbers were only a problem in the 1986-87 sample, with 200 cases (out of the 1,643 that entered this step of the construction of the analytical sample) dropped because we had no access to their post-release arrest information. The number of cases dropped for the 1981-82 and 1991-92 samples was 7 and 26, respectively.

It appears that CII numbers were unavailable for some wards in the 1986-87 sample because their OBITS records had already been purged (due to their prior discharge) when the CII identification numbers for this sample were obtained. This occurred when the sample was initially drawn (Skonovd and Haapanen 2000). We

compared the prior arrest histories of those missing arrest data and those not missing this information and found that the cases missing their CII numbers (and thus missing their post-release arrest data) had a higher mean number of prior arrest charges. The cases missing (n = 200) their post-release data had an average of 11.79 prior arrest charges for criminal offenses, whereas the group (n = 1443) with CII numbers intact averaged 10.22 prior arrest charges. Further, a comparison of parole performance between the two groups indicated a similar finding: 90% of those missing their CII numbers were given a "Dishonorable Discharge" (and only 1 earned an "Honorable Discharge"), whereas 78% of those with valid CII data were "Dishonorably Discharged" (9% of them earned an "Honorable Discharge.").

It is simply impossible to reliably impute a longitudinal pattern of arrest charges (over an extended segment of the age distribution) for cases that were missing either the prior or post-release arrest information. Thus, cases missing either of these portions of their criminal arrest histories were excluded from the final analytic samples. Missing arrest data proved not to be a problem for either the earliest (1981-82) or the latest (1991-92) release samples. For the 1986-87 sample, however, we suspect that, on average, there is little bias that results from missing data because an examination of those in the sample with and without missing arrest data were equally distributed among the highest and lowest parole failure risk offenders. Perhaps more importantly, it bears noting that even the lowest risk cases in these samples still have incredibly high failure rates. Complete arrest information was available in both the pre- and post-arrest periods for over 80% of the male CYA regular cases. After removing the cases missing arrest data in

either the pre- and/or post-release periods, the resulting sample sizes were 1,991 (1981-82), 1,443 (1986-87), and 1,434 (1991-92) respectively.

Two of the males in the 1981-82 sample were reported to have died, but the dates of their deaths were not recorded. Neither of these individuals had any arrests in the post period; these two cases were dropped since it was unknown how long they were on the street before their deaths.

After removing the cases that failed one of the five steps in the hierarchical process, we obtained the final analytic samples that are used in the analyses presented herein.

Data Limitations

Before concluding this section on the nature of the data used in this study, a discussion of the possible limitations of the data is necessary. The first limitation of this study is that both the arrest and mortality data only use records from the state of California. To the degree the wards migrated outside of California and either died and/or were arrested elsewhere, the data will undercount the extent of these outcomes. We do not feel this is a fatal limitation within the serious youthful offender population studied here, however, for two reasons. First, most of these wards (over 95%) show-up in one of the sources of data in the post-release period (i.e., they either died, were arrested, and/or were incarcerated in California at some point during the post-release period). Second, many of these wards were often on parole in the post-release period (either from the CYA or the CDC), and thus most of them had conditions of parole release that prohibited them from leaving the state of California without permission of their parole officers (not that

they always obeyed the conditions of their parole). Nonetheless, it is necessary to keep in mind that these data refer only to records from the state of California and thus may underestimate arrest (and mortality) particularly if they occurred outside California.

A second limitation of this study is that the analyses are based entirely on official criminal justice data (i.e., arrest records). There are no self-report data available for these samples. The strength of self-report data is that it may allow for the investigation of hidden criminal activity patterns that do not depend on the offender being apprehended by law enforcement officials. Official arrest data, on the other hand, are entirely dependent on apprehension, and thus offenders who are actively offending but never get arrested do not appear in the official data records as an offender. To the degree that the wards in this sample were committing criminal acts and were not being arrested for them, the analyses here would understate the extent of their criminal activity. There is little doubt that the individuals in our samples committed many crimes for which they were not arrested, but there are three points that counter the argument that this limitation is a serious impediment to our analyses. First, the majority of studies comparing self-report data to official arrest data find that those offenders who report the most frequent and serious offenses are also consistently the most frequent and serious offenders in the official data (see, e.g., Hindelang et al. 1979, 1981; Huizinga and Elliott 1986; Farrington 1989). As Farrington (1989: 418) notes, self-report and official data generate "comparable and complementary results on such important topics as prevalence, continuity, versatility, and specialization in different types of offenses." Second, self-report data are not without criticism, especially with respect to the topics of interest in this study. Recent criticisms of self-report data include issues surrounding both the

validity of these data (Piquero et al. 2002) and the reliability of using self-report data to examine within-individual changes in the relationship between age and crime (Lauritsen 1998, 1999).

Third, collecting self-report data that would allow one to examine the issues addressed in this study and to make reliable generalizations concerning the population of serious youthful offenders would be both economically and practically infeasible. Recall that the wards of the CYA represent less than 5% of the known (arrested and processed) delinquents *in the entire state* of California. If one only wanted to study issues related to the development of criminal offending patterns of these serious offenders as they age, the initial sample size that would be required to encapsulate a considerable number of serious offenders (that are comparable to the CYA wards) in a sample would have to be so large that the research would be economically infeasible. This is especially true when you take into account that in order to reliably record the self-reported offense patterns at any given age, the interviews, beginning in early childhood, and would have to be conducted annually (across the entire state of California).

Certainly, some researchers would not agree with this conclusion, and believe that serious violent offenders can be studied through self-report data found in samples of the general population. For example, Elliott (1994: 17) has argued that "truly serious violent offenders are included and retained in longitudinal general population studies. In fact, persons with arrest histories and incarceration experience are among the most easily tracked, and seldom are lost in longitudinal studies."[14] Our experience with the CYA

[14] There no doubt is a qualitative divide between what Elliott (1994) calls serious violent offenders and the youthful offenders in the CYA. To make this point clear, consider the following. Elliott (1994: 18) noted that by age 27 there was a 30% cumulative prevalence rate of serious violent offending (i.e., admitting that

data used in the presents study leads us to question this conclusion. The serious youthful offenders in our samples are the hardest of all offenders in the state to track because they are often literally and figuratively "under the gun" of law enforcement officials, and thus often can't be found at home. They are not hard to track in the sense that they have high migration rates across state lines, but this doesn't mean they are easily tracked within their neighborhoods (and many researchers would not dare enter some of these neighborhoods at night or even during the day). Let's suppose that you could easily locate them. These individuals both by their nature and legal status would probably not be open to the intrusive questioning by researchers about the patterns of their criminal activity, even after assurances of anonymity. Many of these individuals are the very same people that had to be contained in metal cages in order to be taught by a CYA teacher (who wears a flak jacket for protection during instruction). These individuals often have problems maintaining scheduled appointments with their parole officers under the threat of a loss of their freedom (and they are often AWOL for periods of time while on parole). Thus, the idea of scheduling a self-report interview and actually obtaining accurate data (that deals with very sensitive information regarding their offending behavior) seems an unrealistic expectation in our opinion. There simply is no better way to study this population than through the use of official data. Thus, one of the suspected limitations of these data (i.e., official data) may actually be its main strength.

---

they had done something seriously violent such as "attacking someone with the idea of seriously hurting them") in the National Youth Survey. In the samples used in this study, by the end of the follow-up 82% of the cases had been *arrested* for a serious violent offense (and averaged 3 serious violent arrest charges such as homicide, aggravated assault, armed robbery, rape, and sodomy); further, roughly 10% of each sample had a least one *homicide* arrest charge in their records.

METHODOLGY

This section describes the analytical methods employed in this study. In this section, we first ignore the specific analytical goals addressed in Chapters 7 and 8 and generally discuss an issue of fundamental importance to this study—modeling a dependent variable that is a count variable in the presence of repeated measurements. At the conclusion of this section, we will return to the specific goals of Chapters 7 & 8 and explicitly discuss the specific analytical approaches to be undertaken in each chapter and the specifications of the regression models employed therein.

The Dependent Variable

The dependent variable assessed in Chapters 7 and 8 is the *count number of criminal arrest charges at each age* for the members of the panels. This count variable does not include arrest charges for probation violations (e.g., program failure, out of control), parole violations (e.g., positive drug test), or traffic offenses (e.g., driving without a license, driving with a suspended/revoked license). Instead, the dependent variable only counts arrest charges regarding the offenses that were more "seriously criminal" in nature (e.g., homicide, robbery, burglary, theft, drug trafficking, possession of a loaded firearm).

In the statistics literature, data like that employed herein are known as *unbalanced panel* datasets because the cases in the samples have varying numbers of records in the final analytic files. The longitudinal offending sequence for each of the wards began at age 7, and the sequence ended with the final age at which the case was known to be "at

risk."[15] For most of the cases, the final age at risk was determined by the end of the follow-up period (June 30, 2000), but for some cases the final age was determined by the age at death. Given that the wards were of varying ages at the time of release from the CYA for the "sample stay" (i.e.. the stay at the CYA that resulted in the ward's inclusion in the sample), the maximum age at which each ward's criminal arrest history was available could also vary substantially. Ages during which the wards were incarcerated in the CYA for the sample stay were removed from consideration of the risk of arrest. Appendix C contains a table with a detailed description of the percentages of each sample that had available criminal arrest histories at each age; here we simply present a brief description of the number of "periods" or "age years" (hereafter referred to as "data points") that were available for analysis in each sample.

For the 1981-82 sample, age 43 was the maximum age at which a respondent's criminal arrest history was available, and roughly 50% of the sample could only be observed through the age of 37. The number of data points used in the panel analyses varied from a minimum of 9 to a maximum of 37, and the average number of data points was 30.

For the 1986-87 sample, the oldest age at risk by the end of the follow-up period was 38, and only 50% of the sample was available for study after the age of 33. The minimum number of data points available for analysis within this sample was 11, the maximum was 32, and the average number of data points was 26.

[15] About 5 cases in each sample experienced their first arrest event at age 6. To keep the absolute size of the datasets to a minimum, those arrests charges were included in the age 7 count for those cases.

The oldest age at which a case's criminal arrest history was available for study in the 1991-92 sample was 33, and the arrest histories for 50% of the sample were only followed through the age 27. The minimum number of data points available for analysis in this sample was 10, the maximum was 27, while the average number was 21.

The dependent variable of this study (arrest counts) has two properties associated with it that must be appropriately taken into account in any statistical model: (1) it is a nonnegative count variable and (2) the data structures contain multiple observations per case (i.e., there is a lack of independence). We deal with each of these issues in turn in the following two sections.

Modeling a Count Variable: The Poisson & Negative Binomial Regression Models

Given that the dependent variable in this study is a nonnegative count variable, the methods of analysis employed here must take into account the discrete nature of this variable.[16] If we were to apply standard OLS linear regression models that assume a continuous, normally distributed dependent variable as opposed to a skewed count dependent variable such as that used here, it would produce biased, inefficient, and inconsistent estimates of the covariates included in the model specification, as well as possibly predicting a negative number of events (King 1988; Long 1997).[17] For these

[16] This discussion of the Poisson model and generalizations of it draw heavily on the detailed treatments of these methods in Hausman et al. (1984), Cameron and Trivedi (1986, 1998), Hardin and Hilbe (2001), and especially Land et al. (1996).

[17] The only case in which this is not true is if the mean rate of event occurrence is large; in that case, the OLS model with Gaussian errors provides a suitable approximation (Land et al. 1996). As the events become more rare (and the mean rate increasingly approaches zero), the normal approximation becomes increasingly less suitable.

reasons, two general regression models based on a probability distribution that explicitly takes into account the discrete nature of count variables have been proposed: (1) the Poisson regression model and (2) the negative binomial regression model.

The Poisson Regression Model

To begin, let us allow y_{it} to denote the observed event count of the ith individual ($i = 1, \ldots, N$) at time (age) t ($t = 7, 8, 9, \ldots, T_i$), where $T_i = \max(\text{age}_i)$. The univariate Poisson probability distribution function is specified as

$$\Pr(Y_{it} = y_{it}) = \frac{e^{-\lambda_{it}} \lambda_{it}^{y_{it}}}{y_{it}!}, \qquad y_{it} = 0, 1, 2, \ldots \tag{1}$$

where $\Pr(Y_{it} = y_{it})$ indicates the probability that the random variable Y_{it} takes on the observed value y_{it} for the ith individual at time t, and λ is the Poisson parameter representing the mean rate of event occurrence at time t (Land et al. 1996).[18] The Poisson distribution, which is a one-parameter probability distribution, makes a strong assumption regarding the relationship between the mean and variance of the random variable Y_{it}. This assumption, known as the *equidispersion* assumption, assumes that the mean and the variance are equal:

$$E(Y_{it}) = \text{var}(Y_{it}) \equiv \lambda_{it}. \tag{2}$$

[18] See Appendix A in King (1988) for a mathematical proof that all event count data that meet a few modest assumptions about the data generation process can be shown to arise from a Poisson process.

More generally, the expected value of a count variable such as this could be written as

$$E(Y_{it}) = \mathrm{var}(Y_{it}) \equiv \lambda_{it}\sigma^2, \tag{3}$$

where σ^2 is the *dispersion parameter* constrained to be greater than or equal to zero (King 1989). A count variable is said to exhibit equidispersion and to be Poisson-distributed if $\sigma^2 = 1$, but if $0 \leq \sigma^2 < 1$ then the variable is said to be *underdispersed*, and if $\sigma^2 > 1$ then the variable is said to *overdispersed* (King 1989; Land et al. 1996; Lindsey 1993, 1995). Overdispersion is very common property among dependent variables utilized in social science data, whereas underdispersion is relatively rare (King 1989).

The important point to make here is that overdispersion implies a significant substantive fact critically relevant to this study: there is unexplained variation in accounting for why some subjects have greater or fewer total arrests (events) than do other subjects. Stated differently, there is more heterogeneity in the mean event rate *among the individuals* than would be expected according to the Poisson distribution. One possible way to account for why some individuals have a higher mean arrest rate than others is to specify a Poisson regression model whereby a set of measured covariates are included through the equality

$$\ln(\lambda_{it}) = \mathbf{X}_{it}\boldsymbol{\beta}, \tag{4}$$

or equivalently

$$\ln(\lambda_{it}) = \mathbf{X}_{it}\boldsymbol{\beta}, \tag{5}$$

where $\mathbf{X}_{it}$ is a matrix of measured covariates on individual i at time t, and $\boldsymbol{\beta}$ is a column vector of regression coefficients relating the covariates to the mean arrest rate.[19]

According to Land et al. (1996), inclusion of measured covariates in the model specification now leads to a conditional expectation function whereby the expected mean and variance of the event count are conditional on the $\mathbf{X}$ matrix such that

$$E(Y_{it}|\mathbf{X}_{it}) = \mathrm{var}(Y_{it}|\mathbf{X}_{it}) = \lambda_{it} = \exp(\mathbf{X}_{it}\boldsymbol{\beta}). \tag{6}$$

Similar to the deterministic relationship stated above in equation (2), equation (6) still implies a deterministic relationship, only now it is a conditional deterministic relationship (conditional on the measured covariates) whereas before it was an unconditional deterministic relationship (Hausman et al. 1984).[20] However, conditional on the observed covariates, the observed relationship is still nonstochastic.

As noted by Land et al. (1996), the equidispersion assumption of the Poisson regression model is an unrealistic expectation in many social science data sets, and furthermore the failure to satisfy the equidispersion assumption leads to underestimated

[19] The logarithmic *link function* is used to link the linear systematic component, denoted as $\mathbf{X}\boldsymbol{\beta}$, to the response variable (Nelder and Wedderburn 1972; McCullagh and Nelder 1989; Hardin and Hilbe 2001) in order to ensure that that event rate is predicted to be nonnegative (Land et al. 1996).
[20] Hardin and Hilbe (2001: 128) show the mean (first derivative) and variance (second derivative) functions of the Poisson distribution are identical.

standard errors and inflated t-ratio tests of significance (see also Dean 1998; Hardin and Hilbe 2001). In other words, applying a Poisson regression model to data that cannot satisfy the equidispersion assumption can cause a covariate to appear to be a significant predictor of the outcome variable, when in fact it is not. In this case, the statistical significance is spurious, due to the consequences of overdispersion. For this reason, methods that "scale" the relationship between the mean and variance were sought (Hardin and Hilbe 2001). The primary method that is used in the presence of significant overdispersion in a Poisson regression model is to estimate a negative binomial regression model.

The Negative Binomial Regression Model

As stated in Chapter 3, it is unrealistic to assume that every factor related to a dependent variable will be measured and included in all datasets, and thus there will always be some inherent variability (e.g., overdispersion) in the event counts between individuals that must be accounted for (Lindsey 1993, 1995; Land et al. 1996). In the absence of the measured covariates that explain the discrepancy, this variation is usually accounted for as *stochastic* or *random variation* in the dependent variable. Indeed, the precise reasoning behind fitting a negative binomial regression model (instead of a Poisson regression model) is to include stochastic variation in the event count (Hausman

et al. 1984; Cameron and Trivedi 1986; Land et al. 1996).[21] The negative binomial regression model is specified as

$$\ln(\lambda_{it}) = \mathbf{X}_{it}\boldsymbol{\beta} + \varepsilon_{it}, \tag{7}$$

or

$$\lambda_{it} = \exp(\mathbf{X}_{it}\boldsymbol{\beta}) \bullet \exp(\varepsilon_{it}), \tag{8}$$

where $\exp(\varepsilon)$ is distributed as $\Gamma(1,\alpha)$. The α parameter is known as the dispersion parameter and plays a defining role in scaling the relationship between the mean and the variance as shown in equation (9) below. The inclusion of the gamma distributed error term allows for unexplained variation in $\ln(\lambda_{it})$ (Land et al. 1996). This unexplained variation can be thought of as having been produced in one of two ways: (1) through the effects of an omitted exogenous variable(s) (Gourieroux et al. 1984a, 1984b) or (2) through inherent stochastic variation (Hausman et al. 1984). The negative binomial model is known in the statistical literature as a *parametric mixed Poisson* regression

[21] The gamma distribution is the "conjugate distribution" for the Poisson distribution, which allows for a closed form solution that is analytically tractable (Lindsey 1995). Assuming the heterogeneity is normally distributed does not lead to an analytically tractable solution (Land et al. 1996; Cameron and Trivedi 1998). Further, Hardin and Hilbe (2001: 142-146) show that the negative binomial regression model can be thought of in one of two ways: (1) as a Poisson model with gamma distributed heterogeneity (where the gamma distribution is constrained to have a mean of 1); or (2) as a regression model based on the negative binomial probability function that is independent of the Poisson model. Regardless of which way one chooses to think about the model, the resulting likelihood functions (as shown in their derivations of the likelihood function for each version) are in fact, identical.

model because a parametric mixing distribution (i.e., the gamma distribution) has been incorporated into the Poisson model.

There are two general formulations of the negative binomial model (Cameron and Trivedi 1986, 1998; Hardin and Hilbe 2001): (1) NB1 that specifies a linear mean-variance relationship; and (2) NB2 that specifies a quadratic mean-variance relationship. These models both consider the variance as a function of the mean (or Poisson parameter) such that

$$\mathrm{var}\left(Y_{it}\middle|\mathbf{X}_{it}\right) = \lambda_{it} + \alpha\lambda_{it}^{p}, \tag{9}$$

where p is equal to 1 in the NB1 formulation of the negative binomial models and p is equal to 2 in the NB2 formulation. The NB1 and NB2 models are both Poisson-gamma mixture models, but each model provides a different specification of the mean-variance relationship. The log-likelihood functions for the NB1 and NB2 models are presented in Cameron and Trivedi (1998) and Hardin and Hilbe (2001).

NB2 is the more commonly used negative binomial model because it is the model that was first programmed into software packages such as LIMDEP (Greene 1998) and Stata (StataCorp 2001) that were commonly used to model count data. As demonstrated by Land et al. (1996), under the assumption that $\exp(\varepsilon)$ is distributed as $\Gamma(1,\alpha)$ and that ε is independent of X (which allows the marginal density of Y_{it} to be derived by integrating with respect to ε), the probability of observing the count y_{it} for the ith individual at time t in the NB2 model is:

$$\Pr(Y_{it} = y_{it}) = \int_0^\infty \Pr[Y_{it} = y_{it}, \lambda_{it}] f(\lambda_{it}) d\lambda_{it} \tag{10}$$

$$= \left(\frac{\Gamma(y_{it} + \nu)}{y_{it}!\Gamma(\nu)}\right)\left(\frac{\nu}{\nu + \lambda_{it}}\right)^{\nu}\left(\frac{\lambda_{it}}{\nu + \lambda_{it}}\right)^{y_{it}}, \tag{11}$$

where $\lambda_{it} = \exp(\mathbf{X}_{it}\boldsymbol{\beta})$, $\Gamma(\cdot)$ is the gamma distribution, $\nu = 1/\alpha$, and $\alpha \geq 0$. Estimates of α and $\boldsymbol{\beta}$ are obtained using maximum likelihood methods (Hausman et al. 1984; Cameron and Trivedi 1998; Land et al. 1996). Under this specification, the mean and variance are (Cameron and Trivedi 1986, 1998; Land et al. 1996):

$$E(Y_{it}|\mathbf{X}_{it}) = \exp(\mathbf{X}_{it}\boldsymbol{\beta}) = \lambda_{it} \tag{12}$$

and

$$\mathrm{var}(Y_{it}|\mathbf{X}_{it}) = \lambda_{it}(1 + \alpha\lambda_{it}) = \lambda_{it} + \alpha\lambda_{it}^2. \tag{13}$$

Thus, the expected number of events is still equal to λ_{it} (or $\exp(\mathbf{X}_{it}\boldsymbol{\beta})$), but the variance is no longer constrained to be equal to the mean; there is now a quadratic relationship between the mean and variance.

Under the NB1 model, a change is made to equation (11): $v = \lambda/\alpha$ instead of $v = 1/\alpha$ (Cameron and Trivedi 1998; Long 1997). The mean and variance under the NB1 specification are:

$$E(Y_{it}|\mathbf{X}_{it}) = \exp(\mathbf{X}_{it}\boldsymbol{\beta}) = \lambda_{it} \tag{14}$$

and

$$\mathrm{var}(Y_{it}|\mathbf{X}_{it}) = \lambda_{it}(1 + \alpha\lambda_{it}) = \lambda_{it} + \alpha\lambda_{it}. \tag{15}$$

Thus, the expected number of events is still equal to λ_{it}, but the variance is now linearly related to the mean through the dispersion parameter.

Cameron and Trivedi (1998) recommend choosing between the NB1 and NB2 model on the basis of the log-likelihood values. The model with the larger (less negative) log-likelihood value is favored since they both are estimated using the same number of parameters.[22] Substantively, however, the importance of the negative binomial model (whether specified as NB1 or NB2) is that individuals with identical values on the included covariates now have gamma distributed expected event counts, rather than being equal to the same conditional mean rate (as in the Poisson model).

[22] The negative binomial models presented in Chapter 8 were estimated using both the NB1 and NB2 specifications. The NB1 specifications always had larger log-likelihood values, and thus the NB1 versions are the ones presented in Chapter 8. It should be noted, however, that the NB2 models generated identical substantive conclusions to those reached with the NB1 models.

In fact, the Poisson model is nested in the negative binomial model (Cameron and Trivedi 1996, 1998; Land et al. 1996; Long 1997; Hardin and Hilbe 2001). The boundary or limiting case corresponds to $\alpha = 0$, under which the negative binomial model becomes the Poisson model. For example, if $\alpha = 0$ in equation (9), then we arrive back at the initial equidispersion assumption of the Poisson regression model found in equation (6). As Land et al. (1996: 398) note, "this circumstance corresponds to the limiting case where all individuals have the same λ_i, conditional on $\mathbf{X}_i$, which is precisely the assumption of the Poisson regression model."

Of course, as α increases in size, so does the overdispersion of the data (Hardin and Hilbe 2001), and thus testing for the presence of significant overdispersion often becomes a primary task when modeling count data. The standard statistical test for assessing overdispersion involves testing the null hypothesis $H_N : \alpha = 0$ against its alternative, $H_A : \alpha > 0$. Because the Poisson model is a nested version of the negative binomial model, the test for significant overdispersion is frequently accomplished via a likelihood ratio test that compares the log-likelihood values of the negative binomial regression and Poisson regression models. The likelihood ratio test statistic is calculated as twice the difference in likelihood values, and this test statistic is distributed as χ^2 with one degree of freedom. While this is how the test for overdispersion has been calculated in the past, currently it is recognized that this form of the likelihood ratio test is incorrect in this particular situation. More specifically, because the dispersion parameter has to be greater than or equal to 0, the null hypothesis sits on the boundary of the parameter space (Self and Liang 1987; Gutierrez et al. 2001). Because the null hypothesis is on the

boundary of the parameter space, a critical regularity condition is violated—"the null parameter space is no longer interior to the full parameter space, and thus the result which states that the likelihood ratio test statistic tends towards χ_1^2 [chi-square with one degree of freedom] in distribution is untrue" (Gutierrez et al. 2001: 16). As shown by Self and Liang (1987), the correct test statistic is a 50:50 mixture (χ_{01}^2) of (1) a chi-square distribution with a point mass at zero (χ_0^2) and (2) a chi-square distribution with 1 degree of freedom (χ_1^2). P-values calculated according to this 50:50 χ_{01}^2 mixture corresponds, in fact, to one-half the p-value calculated using only the upper tail area of the chi-square distribution with 1 degree of freedom.[23]

Land et al. (1996) note that the negative binomial regression model is a significant generalization of the Poisson model because it accommodates overdispersion in count data while simultaneously keeping the suitable features necessary to model count data. However, Land et al. (1996) also note that the negative binomial regression is also a restrictive model because it assumes that the heterogeneity is gamma distributed in the population, which maybe an arbitrary assumption. If this assumption is incorrect, the standard errors of the regression coefficients will be spuriously deflated leading to inflated t-ratios.

More importantly, however, the negative binomial regression model, in the most basic form (as specified in this section), completely ignores the dependence among observations when it is applied to panel data. This is significant because serial dependence among observations is known to be one of the most significant causes of

[23] See the simulations of Gutierrez et al. (2001) for an example.

overdispersion in count data (Lindsey 1993; 1995; Winkelmann 1995, 1997; Dean 1998; Pickles 1998). Recall that overdispersion substantively implies that the model fails to account for why some subjects have greater or fewer total arrest events than other subjects, or stated differently, there is more heterogeneity in the outcome variable *among the individuals* than would be expected according to the probability distribution.

In the presence of panel data, the negative binomial model specified here is often referred to as a "naïve pooling" model (Burton et al. 1998; Hardin and Hilbe 2001) because it *naively* treats the panel data set as a *pooled* sample consisting of N*T independent individuals rather than as simply N independent individuals each with T (or T_i if unbalanced) dependent observations (Hamerle and Ronning 1995). For example, the negative binomial model expressed above treats the extra variation as resulting purely from transient stochastic variation, rather than allowing a main component of the extra variation to be caused *explicitly* by the stochastic dependence between the observations within the N individuals (Dean 1998; Lindsey 1993, 1997; Cameron and Trivedi 1998). Within each individual, each "draw" (i.e., for a given "age" record) of the random effect from the gamma distribution is completely independent of the other draws for that individual (i.e., for the other "age" records of that individual). Stated more emphatically, the standard version of the negative binomial model (that ignores the panel structure of the data) does not control for persistent unobserved heterogeneity, or as Hsiao (1986: 158) notes, "statistical models developed for analyzing cross-sectional data essentially ignore individual differences." The stochastic variation of the event counts for each individual's panel records is viewed as having been generated by chance—there is no serial dependence of the individual's records. "A shortcoming of the negative binomial

model is that it does not allow for firm [individual] specific effects so that serial correlation of the residuals (i.e., nonindependence of the counts) may be a problem" (Hausman et al. 1984: 922).

This failure to correct for the dependence among the observations is particularly critical for the topic of this study because the "naïve pooled" negative binomial model completely ignores two possible sources of overdispersion in the data—population heterogeneity and state dependence—because it ignores the serial dependence within the data (Hsiao 1986; Lindsey 1993, 1995; Cameron and Trivedi 1998; Dean 1998). For example, if "individuals who have experienced an event in the past are more likely to experience the event in the future than are individuals who have not previously experienced the event," then this will induce overdispersion in the data (Heckman 1981b: 91). In such cases the key question is whether this overdispersion (which is a consequence of the serial dependence) is the result of a process of contagion/state dependence, population heterogeneity, or possibly both of these processes.[24] These two sources of overdispersion are the same two explanations discussed in Chapter 2 as fundamentally critical to our study (because they are rival hypotheses in explaining the relationship between past and future criminal activity). As Hamerle and Ronning (1995: 411-412) state, ignoring "heterogeneity among cross-sections or time-series units [individuals] could lead to inconsistent or meaningless estimates of the structural

[24] This issue has been raised not only in studies of criminal behavior, but also in studies of accidents (Bates and Neyman 1952; Greenwood and Yule 1920), unemployment (Heckman 1981b; Heckman and Borjas 1980), bouts of schizophrenia (Kessing et al. 1999) and emotional distress (Fischer et al. 1984; Robins 1966, 1978), and Medicare claims for Alzheimer's/dementia among the elderly (Taylor, Fillenbaum, and Ezell 2002).

parameters...controlling for heterogeneity is in most applications a means to obtain consistent estimates of the systematic part of the model."

For example, suppose in the pooled negative binomial model one were to find a significant association between a binary indicator of arrest at the previous age and arrest at the current age. In the standard negative binomial model, the process underlying this significant association would be indeterminable because the effects of persistent individual differences are left uncontrolled in this model. Lindsey (1993: 157) pointedly remarks, "if a missing variable [underlying criminal propensity] can be assumed constant over all events on a unit, but differs among units, this will yield stochastic dependence among the events on each unit," and this missing variable will, in fact, "induce an effect identical to apparent contagion" or state dependence.

It has been shown, however, that the unique structure of panel data can be exploited to investigate the above two critical sources of serial correlation of an outcome variable across waves or periods (Heckman 1981a; Hsiao 1986; Hamerle and Ronning 1995; Cameron and Trivedi 1998; Powers and Xie 2000). For example, in an early study investigating whether population heterogeneity or state dependence was driving overdispersion in accident data, Neyman (1965: 6) noted that the distinction between these two processes would be possible if "one has at one's disposal data on accidents incurred by each individual separately for two periods of six months each." Thus, with more than two waves or periods of data on a set of individuals, models can be estimated that attempt to disentangle the effects of population heterogeneity from those of state dependence by specifically incorporating sources of "hidden" or unobserved heterogeneity.

Accounting for Serial Dependence: Persistent Individual Differences

In this section, we discuss the two most common methods that are used to control for persistent individual differences in panel data: parametric random effects models and semiparametric random effects models.[25] Before presenting the technical aspects of each formulation, we first broadly compare the two different methods on the basis of how each model accounts for persistent unobserved heterogeneity.

In the parametric random effects model, the error term is specified to be composed of two components (Heckman 1981a; Hsiao 1986; Nagin and Paternoster 1991; Hamerle and Ronning 1995):

$$\varepsilon_{it} = \alpha_i + u_{it}, \tag{16}$$

where α_i represents a persistent (time-stable) individual-specific component that is assumed to follow a specific parametric distribution and u_{it} is a stochastic component that follows some specified parametric distribution. Thus, the parametric random effects model assumes that the persistent unobserved heterogeneity follows a known, mathematically tractable parametric distribution that is specified by the user.

The semiparametric random effects model, on the other hand, makes no parametric assumption about the distribution of unobserved heterogeneity, but rather this

[25] Fixed effects estimators are not included here or employed in this study because fixed effects models with lagged values of the outcome variable prohibit inclusion of time trends or age effects (Hamerle and Ronning 1995; Cameron and Trivedi 1998; Bushway et al. 1999). Due to the incredibly strong relationship between age and crime (see Chapter 7), these models are clearly inappropriate for modeling crime data and therefore are not considered here.

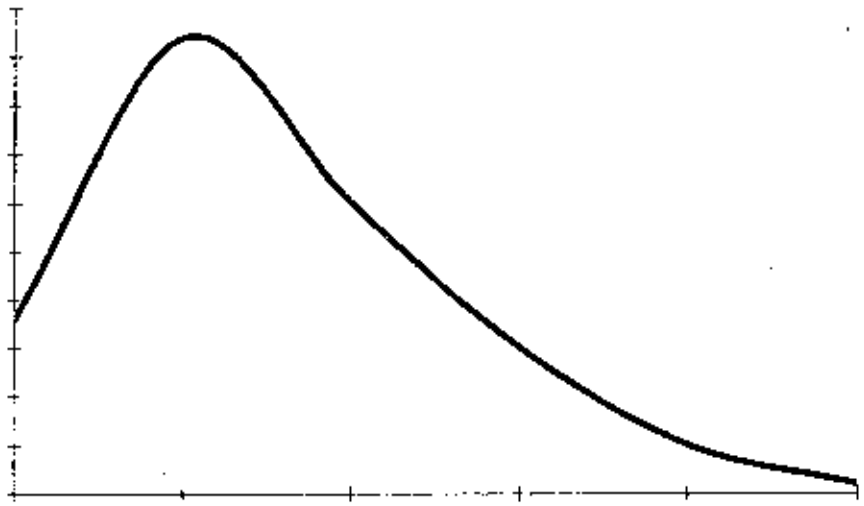
method nonparametrically approximates the distribution of persistent unobserved heterogeneity via set of discrete "points of support." The method only assumes that the distribution of unobserved heterogeneity can be approximated by a discrete multinomial probability distribution (Heckman and Singer 1984; Nagin and Land 1993; Land et al. 1996; Nagin 1999).

Figure 5.2 presents a graphical depiction of the differences between these two models. Panel A of Figure 5.2 represents a continuous mixing distribution (resembling a gamma or beta distribution), whereas Panel B indicates how a continuous distribution can be approximated by a discrete, multinomial distribution using a finite number of "points of support." Panel B contains the same distribution as in Panel A, only in this panel we have used 5 "points of support" (the black "columns" in Panel B) to approximate the continuous distribution. Alternatively and equivalently, the points of support could also be viewed as the histogram "bins" propping-up the continuous distribution.

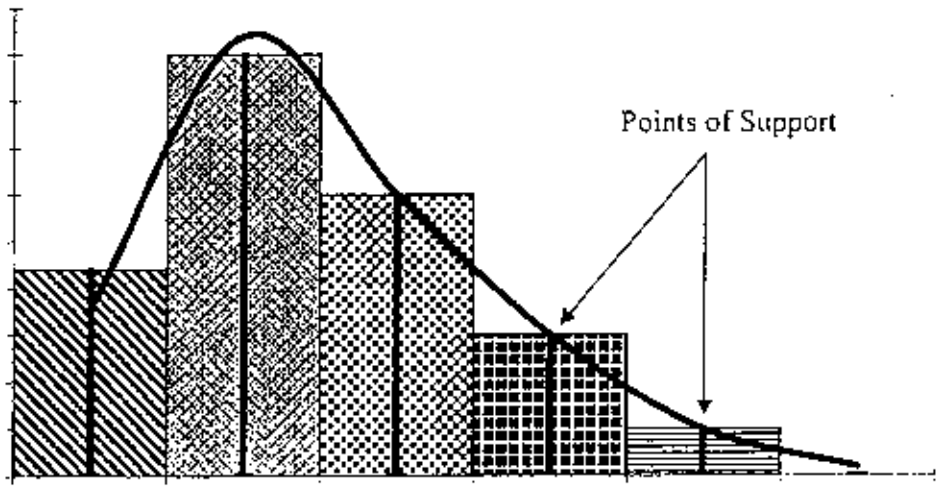
The distinction between these two methods can be viewed in light of the tension in statistics that is ever-present between "parametric" and "non-parametric" methods (Bushway et al. 1999). Parametric methods are more restrictive methods, but if the parametric assumption is appropriate *in the population*, then this method of estimation will be more statistically efficient (i.e., it will have less variance from sample to sample). The non-parametric methods are less efficient if the true distribution is a (mathematically tractable) known continuous distribution, but since these methods do not assume that the mixing distribution follows a restrictive mathematical parametric form *a priori*, they can approximate any continuous distribution regardless of its shape.

Figure 5.2. Approximating a Continuous Distribution with a Finite Number of "Points of Support"

Panel A



Panel B



As Nagin and Tremblay (1999: 1188) note, "the cost of approximation is obvious. Approximations are just that—there is a loss of accuracy. Balanced against this are gains in generality and flexibility. Generally we have no empirical or theoretical basis for specifying the distribution of the growth curve parameters [random effects] within the population." The choice of a parametric mixing distribution is generally made purely on the fact that the some distributions (e.g., conjugate distributions) make the model more mathematically tractable because they ensure that the marginal density of such models have a closed form solution (Cameron and Trivedi 1998). For example, in the standard (single record per individual) Poisson model, the gamma distribution is the conjugate distribution that allows for a closed form solution to the negative binomial regression model. Assuming that the heterogeneity is normally distributed in the standard Poisson model leads to mathematically intractable models. Although in some situations the available mathematically tractable mixing distribution makes substantive sense, in other cases this is unlikely to be true. Indeed, this was the precise reasoning of the thought behind the development of finite mixture models: a particular mixture distribution does not have to be used simply because it is mathematically tractable. The discrete mixture methods allow the data to *speak for themselves* with respect to the nature and extent of unobserved heterogeneity.

Parametric Random Effects Negative Binomial Model

The parametric random effects specification of the negative binomial model was first presented by Hausman et al. (1984), who specified it as

(17)

$$\ln(\lambda_{it}) = \mathbf{X}_{it}\boldsymbol{\beta} + \varepsilon_{it}, \qquad \varepsilon_{it} = \alpha_i + u_{it}.$$

This specification differs from the negative binomial specification of equation (7) in the decomposition of the error term, which in the standard model is specified as $\varepsilon_{it} = u_{it}$. In the random effects formulation of the negative binomial model, the decomposition of the error term results in one component, α_i, representing the fixed, individual-specific component, and one component, u_{it}, representing the transitory stochastic variation. Substantively this model allows for randomness both between-individuals and within-individual across time (or age) (Hausman et al. 1984: 927).[26] The random effects negative binomial model yields a negative binomial model for the ith group that has constant dispersion within the ith group, but the dispersion varies randomly between groups. According to this model,

$$\Pr\left(Y_{it} = y_{it} \middle| \delta_i\right) = \left(\frac{\Gamma(\lambda_{it} + y_{it})}{\Gamma(\lambda_{it})\,\Gamma(y_{it}+1)}\right)\left(\frac{1}{1+\delta_i}\right)^{\lambda_{it}}\left(\frac{\delta_i}{1+\delta_i}\right)^{y_{it}}. \qquad (18)$$

Further, in this the model the ratio $1/(1+\delta_i)$ is assumed to be randomly distributed according to the beta distribution, with the r and s parameters of the beta distribution,

[26] The Poisson random effects model for panel data, which generally assumes gamma distributed heterogeneity, only allows for the individual-specific component (which accounts for between-individual differences) (Hausman et al. 1984; Hamerle and Ronning 1995; Cameron and Trivedi 1998). Conditional on the covariate vector, an individual's expected rate does not have variation over the panel because the error component of the Poisson random effects model is entirely composed of the individual-specific (α) component. Conversely, the negative binomial random effects model allows the mean rate to vary across time because each year is a realization of the gamma probability distribution (Hausman et al. 1984)

$B(r,s)$, estimated from the data at-hand.[27] This results in the following joint probability of events counts for the ith group (Hausman et al. 1984):

$$\Pr\left(Y_{i1} = y_{i1}, Y_{i2} = y_{i2}, \ldots, Y_{in_i} = y_{in_i}\right) \tag{19}$$

$$= \int_0^\infty \prod_{t=1}^{n_i} pr\left[Y_{it} = y_{it} \middle| \delta_i\right] f(\delta_i)\, d\delta_i \tag{20}$$

$$= \left(\frac{\Gamma(r+s)\Gamma\left(r + \sum_{t=1}^{n_i} \lambda_{it}\right)\Gamma\left(s + \sum_{t=1}^{n_i} y_{it}\right)}{\Gamma(r)\Gamma(s)\Gamma\left(r + s + \sum_{t=1}^{n_i} \lambda_{it} + \sum_{t=1}^{n_i} y_{it}\right)} \right) \prod_{t=1}^{n_i} \frac{\Gamma(\lambda_{it} + y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it}+1)}, \tag{21}$$

where $n_i = T_i$ or $\max(t)$ for the ith individual. The resulting log-likelihood for equation (21) is

$$\begin{aligned} \ln L = {} & \sum_{i=1}^{n} \ln\Gamma(r+s) + \ln\Gamma\left(r + \sum_{t=1}^{n_i} \lambda_{it}\right) + \ln\Gamma\left(s + \sum_{t=1}^{n_i} y_{it}\right) - \ln\Gamma(r) - \ln(s) \\ & - \ln\Gamma\left(r + s + \sum_{t=1}^{n_i} \lambda_{it} + \sum_{t=1}^{n_i} y_{it}\right) + \sum_{t=1}^{n_i} \left[\ln\Gamma(\lambda_{it} + y_{it}) - \ln\Gamma(\lambda_{it}) - \ln(y_{it}+1)\right] \end{aligned} \tag{22}$$

The beta distribution is a flexible distribution because it has two parameters (Greene 2000), but it should be remembered that the beta distribution is used in the negative binomial random effects model precisely because it produces a mathematically tractable expression that allows the unobserved random effects to be integrated out without encountering serious numerical complications. Instead of assuming that the

[27] This model was estimated using Stata Version 7 (StataCorp 2001). The distribution of dispersion (noted here) programmed into Stata is the inverse of the Hausman et al. (1984) method, which is just a technical preference of StataCorp. Regardless of the whether δ is estimated using the parameterization employed by Hausman et al. (1984), $\delta/(1+\delta)$, or the inverse parameterization employed in Stata, $1/(1+\delta)$, the resulting solutions are identical

unobserved heterogeneity is distributed according to the beta distribution, we next discuss the semi-parametric formulation whereby a discrete set of nonparametric "random effects" are used to account for unobserved heterogeneity. Again, the finite mixture method is critically important for this study because it has been shown that some of the results concerning the relationship of past to future criminal behavior may have possibly been methodological artifacts resulting from the parametric specification of the random effects. And, further, this method also allows us to investigate the nature of the age-crime relationship within latent classes of serious youthful offenders.

The Semiparametric Mixed Poisson Model

Before describing the technical aspects of the semiparametric mixed Poisson model, we first present a non-technical discussion of the semiparametric mixed Poisson model.[28] In brief, the model assumes that the distribution of unobserved heterogeneity can be "segmented" into a finite number of discrete groups—each of the groups are internally homogeneous with respect to the nature of the unobserved heterogeneity within the group, but there is significant heterogeneity between the groups. According to Nagin and Land (1993) (see also Land et al. 1996), the simplest specification of this model is

$$\ln\left(\lambda_{it}^{j}\right) = \left(\beta_0 + \bar{\varepsilon}_j\right) + \mathbf{X}_{it}\boldsymbol{\beta}, \qquad (23)$$

[28] It should be noted here that while this discussion centers on the Poisson version of the finite mixture model, the finite mixture model is a general class of models that extends far beyond the formulation of the model on the basis of the Poisson distribution. Finite mixture models can be estimated for any distribution in the exponential or multivariate exponential family (e.g., binomial, normal, multinomial logit, and censored normal probability distributions) (see Wedel and Kamakura 1998; Nagin 1999; and Vermunt and Magidson 2001).

where β_0 is the overall constant of the model, $\bar{\varepsilon}_j$ is a constant term that is specific to the jth discrete group or latent class ($j = 1, 2, \ldots, J$), $\mathbf{X}_{it}$ is a matrix of measured covariates on individual i at time t, and β is a column vector of regression coefficients.[29] Cameron and Trivedi (1998: 129) refer to this model as a *random intercept model* because each latent class has a separate constant or intercept parameter assigned to it. The effects of the regression coefficients are constrained to be equal across the groups, and the latent classes differ only with respect to their "location parameter." Nagin and Land (1993) describe this model as producing "constant shifts" in the mean rate. That is, the trajectories of each class are identical in shape, but they differ in the mean location of the trajectory. It bears noting that this is the precise specification that corresponds to Gottfredson and Hirschi's hypothesis concerning the robust nature of the relationship between age and crime—groups differ on their mean offending rate, but the actual shapes of the trajectories are identical.

The bare essence of this finite mixture model is that the finite number of intercept coefficients—there is a separate intercept coefficient for each group or "point of support"—represent the "average contribution" of persistent unobserved heterogeneity to the expected Poisson rate for "individuals possessing levels of unobserved heterogeneity in the immediate vicinity of the jth point of support" (Nagin and Land 1993: 338). This model was subsequently referred to as "semiparametric" in nature because it "combines a parametric specification of the regression component of the model with a non-parametric specification of the error term" (Land and Nagin 1996: 170).

[29] The latent classes or groups are also commonly referred to as the mixture "components" (Cameron and Trivedi 1998) and "segments" (Wedel and Kamakura 1998)

Nagin and Land (1993) also present a more general form of the mixture model

$$\ln\left(\lambda_{it}^{j}\right) = \left(\beta_0 + \tilde{\varepsilon}_j\right) + \mathbf{X}_{it}\beta^{j} \tag{24}$$

where β^{j} is a vector of group-specific regression coefficients. By permitting the regression coefficients to vary across the latent classes, this model allows for full heterogeneity not just in the "location parameter," but also in the nature of each latent class' offending trajectory over time. It is also possible for some of the regression coefficients to be constrained so that they are equal across the latent classes, while also simultaneously allowing other coefficients to vary across the latent classes. Wedel et al. (1993) refer to this type of model as a model with random effects in both the intercept and slope parameters. For example, consider the case where the **X** matrix contains two variables: age and age-squared. By permitting the age coefficients (i.e., growth parameters) to vary across the latent classes, this specification of the mixed model not only allows for heterogeneity in the mean event rate at a given time t, but also for the developmental shape of the trajectories (Nagin 1999; Nagin and Tremblay 1999).

Before concluding this discussion of these finite mixture models, it may be helpful to discuss briefly the more technical side of these models.[30] Let us begin by denoting individual i's longitudinal offending sequence as the vector $\mathbf{Y}_i = \left[y_{i1}, y_{i2}, \ldots, y_{iT_i}\right]$, and allow m_j to denote a random variable indicating the

[30] Readers interested in further details than presented here are referred to Nagin and Land (1993), Wedel et al. (1993), Land et al. (1996), Land and Nagin (1996), Roeder et al (1998), Wedel and Kamakura (1998), and Nagin (1999)

proportion of the dataset estimated to belong to the jth point of support. The random variable m_j is postulated as a draw from a "super-population"—the super population is an additive "mixture" of J discrete populations (Cameron and Trivedi 1998). It is important to note that all of the model parameters in the finite mixture model are jointly estimated, including the proportion of the dataset that is estimated to belong to the jth point of support. The estimated proportion belonging to each latent class is calculated using the logit function

$$m_j = \frac{\exp(\Theta_j)}{\sum_j [\exp(\Theta_j)]}, \tag{25}$$

where the following constraints are is imposed: $m^j \geq 0$ and $\sum_j m^j = 1$. The probability of observing y_{it} events for individual i in time period t in group j is

$$f(y_{it}; \lambda_j), \tag{26}$$

where $f(\bullet)$ is the Poisson density function, and the probability of the entire sequence of individual i at the jth point of support is

$$P^j(Y_{it} | \Theta^j) = \prod_t f(y_{it}; \lambda_j). \tag{27}$$

Now the unconditional probability of observing individual i's longitudinal sequence can be calculated by aggregating the likelihood function (i.e., aggregating the conditional likelihoods) for individual i over the j points of support according to

$$P(Y_i) = \sum_j m^j \bullet P^j\left(Y_i \middle| \Theta^j\right), \quad m^j \geq 0, \quad \sum_j m^j = 1, \tag{28}$$

which is simply the probability of Y_i at the each of the jth point of supports multiplied by the proportion of the population estimated at that point of support and then summed over the J points of support (Land et al. 1996). The log-likelihood function of this model is

$$\ln L = \sum_i \ln(P(Y_i)) = \sum_i \ln\left(\sum_j m^j \bullet P^j\left(Y_i \middle| \Theta^j\right)\right). \tag{29}$$

Of course, a key issue related to the finite mixture model concerns the number of points of supports/latent classes/groups to include in the mixture (D'Unger et al. 1998).[31] In other words, how does one decide how many points of support to include in the model? Unfortunately, a finite mixture model with $J = 2$ mixture components is not nested in the model with $J = 3$ components, and therefore a likelihood ratio test statistic cannot be used to determine the number of components in the mixture distribution "because there is not a unique way of obtaining the null hypothesis from the alternative hypothesis" (D'Unger et al. 1998; Ghosh and Sen 1985; Land et al. 1996: 424; Nagin

[31] To connect this issue back to the "urn schemes" discussion in Chapter 2, this issue is directly related to the "total number of urns in the population" (Nagin and Paternoster 2000: 120).

1999; Titterington , Smith, and Makov 1985). More specifically, the degrees of freedom cannot be uniquely determined because there is more than one way for a latent class to become superfluous: (1) the proportion of the population in the group can be set to zero or (2) the set of parameters for one group can "collapse" onto another group. Therefore, one is prevented from calculating the appropriate degrees of freedom for the likelihood ratio test.

Given this problem with the likelihood ratio test, alternative methods of determining the number of mixture components have been recommended. For example, D'Unger et al. (1998) propose the use of the Bayesian Information Criterion (BIC) statistic as a basis of selecting the appropriate number of groups in the mixing distribution (see also Nagin 1999). The BIC statistic is calculated as:

$$\text{BIC} = \log(L) - 0.5 * \log(N) * (k) \tag{30}$$

where $\log(L)$ is the model log-likelihood value, N is the sample size, and k is the number of estimated parameters. The BIC statistic rewards parsimony because each additional fitted parameter results in a "penalty" proportional to the log of the sample size (Kass and Raftery 1995; Raftery 1995; Nagin 1999). In short, the BIC statistic follows the principle of "Occam's Razor" and values parsimony. Nonetheless, the substantive goal of the BIC statistic is the same as the likelihood ratio statistic—find the best model with the fewest number of parameters. The model that has the least negative value of the BIC statistic is the favored model (Nagin 1999).

The selection of the optimal number of points of supports for the mixing distribution is complicated by the fact that mixture models are known to often have a problem with "local solutions" (i.e., a unique global solution is not always achieved) (Goodman 1974; Cameron and Trivedi 1998; Wedel and Kamakura 1998; Vermunt and Magidson 2001). This issue concerns whether the likelihood function is unimodal or multimodal—it is possible to for the model's algorithm to converge to a local maximum, which is not a true global solution. Simulations by Wedel and Kamakura (1998) indicate that the probability of a local solution increases (1) as the number of mixture components increases; (2) as the number of parameters estimated increases (e.g., including a large number of covariates); (3) when the mixture components are not well separated, resulting in *weak identification* of the model (i.e., the model is "overparameterized" and the estimated groups are not that dissimilar); and (4) when using the Poisson and binomial probability distributions. Several authors have recommend testing for the presence of local solutions through the use of "multiple sets of starting values" in order to test for the presence of local solutions (see e.g., Wedel and Kamakura 1998; Cameron and Trivedi 1998, Greene 2000; Vermunt and Magison 2001). In this study, we follow the recommendations of D'Unger et al. (1998) and Nagin (1999) and guide the selection of the optimum number of mixture components using the BIC statistic. We also undertake extensive testing for the presence of local solutions.

While the finite mixture model is a method of controlling for unobserved heterogeneity, depending on the goals of a particular analysis, each individual can be "assigned" to the latent class to which the individual has the highest probability of belonging. Analyses can then be conducted using either the latent class membership

variables (the set of J binary variables indicating whether the individual was assigned to the jth latent class) or within each of the "latent classes" (i.e., using only the individuals assigned to a particular latent class) (Nagin and Paternoster 2000). For example, graphs of the offending trajectories depicting the relationship between age and crime within each of the latent classes could then be computed. Another alternative is to include the set of J binary variables in regression models as variables that control for persistent unobserved heterogeneity.

Thus, often times one of the key steps of the application of finite mixture models is to sort the individuals in the sample into the latent class for which they have the highest probability of belonging. As shown in Nagin (1999), the *posterior probability* of membership in the jth latent class for individual i is calculated as

$$\hat{P}(j|Y_i) = \frac{\hat{P}(Y_i|j)\hat{m}_j}{\sum_j \hat{P}(Y_i|j)\ \hat{m}_j} \tag{31}$$

This probability represents the estimated probability (based on the model coefficients) of observing individual i's longitudinal sequence, $\mathbf{Y}_i$, given membership in the jth latent class, and the estimated proportion of the population in latent class j. These J probabilities are *posterior probabilities* calculated after model estimation based on the maximum likelihood estimates of the model. Having now completed a description of the various statistical models employed in this study, we now turn to a specific discussion of the analytical approaches to be used in Chapters 7 and 8 of this study.

Analytical Approach Employed in Chapter 7

As discussed in Chapter 3, the focus of Chapter 7 concerns the relationship between age and crime among latent classes of serious youthful offenders. In Chapter 7, we first estimate the following semiparametric mixed Poisson model using the final analytic sample constructed for each of the three CYA samples:

$$\ln\left(\lambda_{it}^{j}\right) = \left(\beta_0 + \bar{\varepsilon}_j\right) + \left((age_{it}/10) * \beta_{age}^{j}\right) + \left(\left((age_{it}^2)/100\right) * \beta_{age^2}^{j}\right). \qquad (32)$$

Models allowing up to 8 latent classes will be tested, and the BIC statistic along with empirical testing regarding whether the solutions are global maximums will be used to select the optimal number of latent classes or components in the mixture distribution. Here we will also test the hypothesis of Gottfredson and Hirschi that there is uniformity in the shape of the age-crime distribution by testing the statistical significance of allowing the magnitude of the age parameters to vary across the latent classes. Following the arrival at the optimal number of latent classes, the sample members will then be assigned to the latent class to which they have the highest probability of belonging. At this point, the basic descriptive features of the offending patterns of the latent classes will be discussed (e.g., age at first arrest, mean number of arrest charges), and then the observed and predicted offending trajectories across age will be graphed and compared. Finally, a descriptive analysis will be undertaken to examine what role, if any, adult

incarceration time may have had on the decline in criminal offending among these three groups of serious youthful offenders.[32]

Analytical Approaches Employed in Chapter 8

In Chapter 8, the relationship between past and subsequent criminal activity will be examined. In that chapter we will employ the use of the multimethod approach of Bushway et al. (1999), which is essentially the "compare and contrast strategy" first described by Heckman and Singer (1984). More specifically, we test the robustness of any observed effect of past and subsequent criminal behavior by employing several different methods of analysis. The relationship between past and subsequent criminal activity will be investigated in five stages.

In the first stage we will employ the use of the semiparametric mixed Poisson model and employ a similar specification to the models estimated in Chapter 7, only in this chapter we also include a binary indicator of arrest in the prior period in the specification:

$$\ln\left(\lambda_{it}^{j}\right) = \left(\beta_0 + \bar{\varepsilon}_j\right) + \left((age_{it}/10) * \beta_{age}^{j}\right) + \left(\left(\left(age_{it}^2\right)/100\right) * \beta_{age^2}^{j}\right) + \left(arr_{t-1} * \beta_{arr_{t-1}}\right) \qquad (33)$$

[32] The adult incarceration data concerning the individuals in these three samples were only obtained on July 22, 2002, and hence the inclusion herein is only descriptive. Future work will examine this question more definitively by creating more complex analytical files that essentially remove the individual from being at risk of arrest during periods of incarceration (through the use of a "street time" exposure or offset variable).

Of central concern in this model will be the regression coefficient, $\hat{\beta}_{arr_{t-1}}$, that estimates the state dependence relationship between the binary variable indicating arrest in the prior period (arr_{t-1}) and the mean offending rate.

In the second stage we will employ the use of the parametric random effects negative binomial model and estimate the following model

$$\ln(\lambda_{it}) = \beta_0 + \left((age_{it}/10) * \beta_{age}\right) + \left(\left((age_{it}^2)/100\right) * \beta_{age^2}\right) + \left(arr_{t-1} * \beta_{arr_{t-1}}\right) + \varepsilon_{it}. \qquad (34)$$

Again, the focus here will be on the regression coefficient, $\hat{\beta}_{arr_{t-1}}$, that estimates the state dependence relationship between the binary variable indicating arrest in the prior period (arr_{t-1}) and the mean offending rate. We will compare the estimates of the state dependence effect from this model using the parametric specification of the unobserved heterogeneity to those obtained in the first stage using the nonparametric specification of the random effects. The specification of the models in the first two stages are identical to the specifications used by Bushway et al. (1999), only here we employ formulations of the statistical models that are consistent with a count variable (whereas Bushway et al. used a binary dependent variable and the probit model).

In the third stage, we employ the use of both the random effects negative binomial model and the standard negative binomial model, and estimate these models using the following specification

$$\ln(\lambda_{it}) = \beta_0 + \mathbf{X}_{LCV}\boldsymbol{\beta}_{LCV} + \left((age_{it}/10) * \beta_{age}\right) + \left(\left((age_{it}^2)/100\right) * \beta_{age^2}\right) + \left(arr_{t-1} * \beta_{arr_{t-1}}\right) + \varepsilon_{it}, \quad (35)$$

where $\mathbf{X}_{LCV}$ is a matrix of binary variables indicating latent class membership and $\boldsymbol{\beta}_{LCV}$ is a column vector of regression coefficients for the latent class indicators. The latent class indicator variables employed in this analysis will be carried over from the latent class results presented in Chapter 7. This stage of analysis allows us to address two previously unanswered questions. First, does including the set of binary latent class indicator variables remove any underlying unobserved heterogeneity? Pending the tests indicating a lack of unobserved heterogeneity after including the latent class indicator variables, the standard negative binomial model will then be employed. Second, we ask if the state dependence effects uncovered in stage two are sensitive so as to allow the age parameters to vary over the latent classes? Bushway et al. (1999) found that models that allowed for "time trend" or age effects significantly reduced the effect of the state dependence variable, but the authors did not test to determine if it was also sensitive so as to allow the age parameters to vary over the latent classes. This issue will be addressed in this stage of the analysis using a set of interaction variables between the age variables and the latent class indicator variables

In the fourth stage, we estimate separate standard negative binomial models on each latent class by itself. This will allow us to test the hypothesis of Moffitt (1993) regarding whether the state dependence effects vary over the latent classes, and specifically whether the effect is particularly pronounced in the adolescent peaked group (pending that such a group is identified in the analyses of Chapter 7). The specification

for this stage will be identical to the specification noted above in the second stage of analysis, and will include the following covariates: age, age-squared, and the binary indicator of arrest at the previous age.

In the fifth analytical stage, we estimate models employing both the random effects negative binomial model and the standard negative binomial model, only here we employ specifically the post release arrest data and also include covariates identifying the theoretically relevant characteristics of wards (e.g., gang member, drug abuser) in the model. This allows us to test the sensitivity of the results and ask the following two questions: (1) would the conclusions of this study have been any different had only the post-release arrest data been available; and (2) are there any covariates significantly related to their post-release offending rate? This analytical stage essentially puts this study in the same research context as the study by Paternoster et al. (1997) and should produce some interesting comparisons.

Before moving on to the main results of this study to be presented in Chapters 7 and 8, we note that in Chapter 6 we will present a descriptive summary of the data used in this study. This is important for documenting the basic facts concerning the criminal offending of the samples, including the nature of their criminal offending behaviors (i.e., what types of offenses for which they were arrested, age at first arrest), certain behavioral characteristics of the subjects (e.g., gang membership, drug abuse), and specifics regarding incidents of mortality and adult incarceration among the members of our samples. For background information purposes, we also present a description of the trends in the crime, arrest, and incarceration rates in California from 1960-2000.

CHAPTER 6

DESCRIPTIVE SUMMARY OF DATA

INTRODUCTION

In this chapter, we present a descriptive summary of the data employed in this study. This description is important for documenting basic information regarding characteristics of the samples, including the basic nature of the extent and breadth of their criminal offending (e.g., number of arrest charges, age at first arrest, the types of offenses for which they were arrested, recidivism rates) and other background factors (e.g., gang membership, history of drug abuse). We also describe the mortality and adult incarceration experiences of the three samples. It is important to note at the outset that the goal of this chapter is not to present a thorough explanation of the significant differences in the distributions of the variables among the three samples. Rather, this chapter focuses on describing the distribution of each of the samples on the critical variables identified in Chapter 5.[1] Differences between the samples will be discussed where they are deemed to be important. Before we present a summary of the data for the

[1] The comparison of significant differences between the samples is not the main focus of this chapter for six main reasons. First, the main goal of the chapter is to give the reader a feel for the nature of the individuals in the samples, and not to get bogged down in whether a variable is significantly different among the samples. Second, for the arrest, mortality, and incarceration data, each of the samples had very different "exposure" or "risk" periods. For example, the 1981-82 sample had 10 extra years of "exposure" or "risk" time compared to the 1991-92 sample, and therefore certain variables will be significantly different if for no other reason than a pure consequence of different rates of exposure. Third, given that there are three different samples, significant differences between samples would involve three separate tests of significance (e.g., 1981 versus 1986; 1981 versus 1986; 1986 versus 1991) for each variable of interest (over 70 variables are presented in Tables 6.1 to 6.6), and this would involve interpreting an excessive number of significance tests that are largely irrelevant to the topic of this study. Fourth, we estimate all of the models in Chapters 7 and 8 on each of the samples by themselves. Fifth, we do not have adequate data that allows us to test or definitively explain why there was a change in the distribution of many of the variables discussed in this chapter, and thus our explanations of the differences would be pure speculation for many of the variables. Sixth, and most important, the overall "substantive story" told in each of the samples is very similar.

wards in this study, we first describe the crime, arrest, and incarceration rates in California from 1960 through 2000 in order to provide an historical context from which to view the data discussed herein.

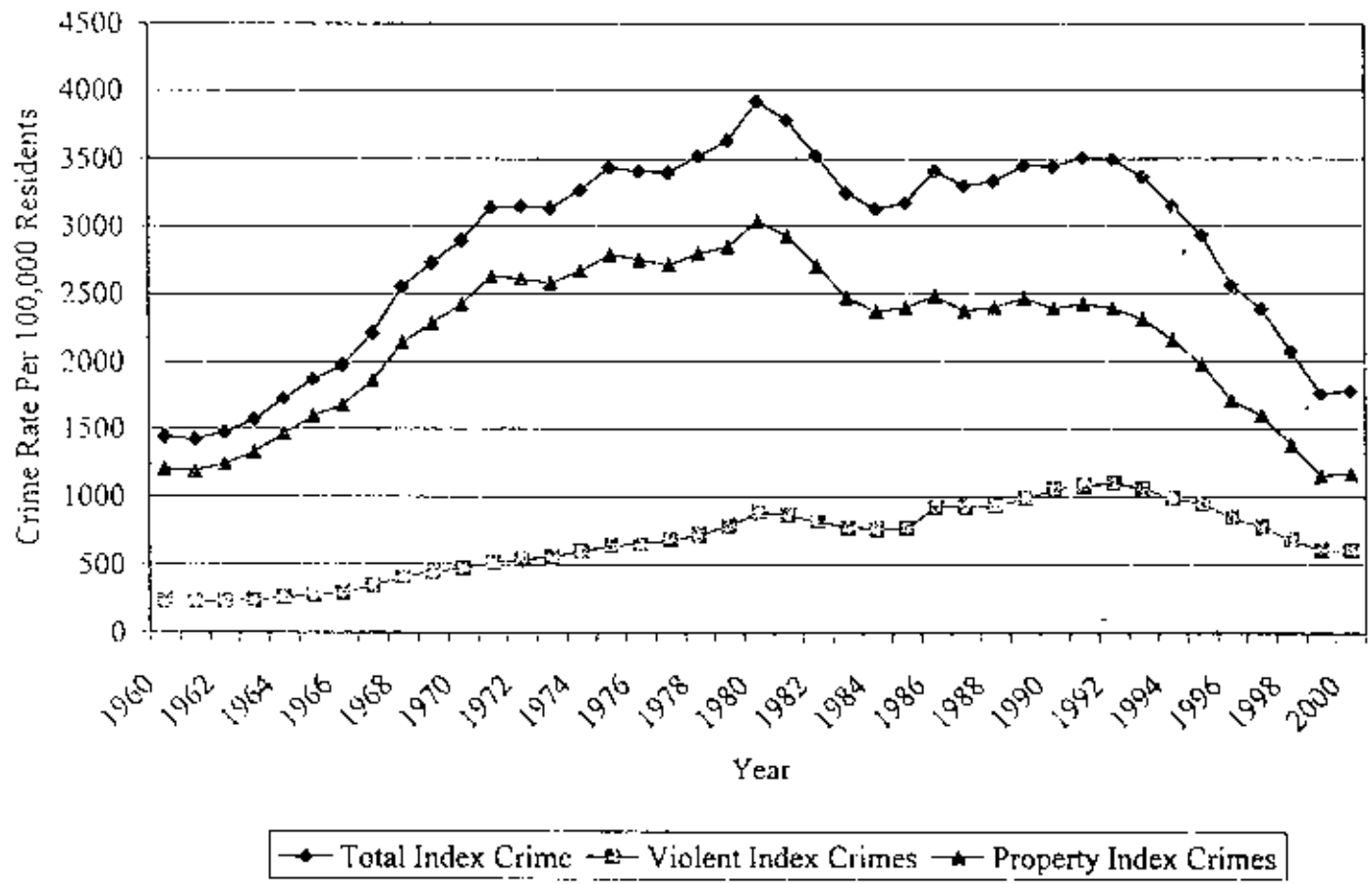
CRIME, ARREST, AND INCARCERARTION RATES IN CALIFORNIA

Three index crime rates in California from 1960 through 2000 are presented in Figure 6.1--the total, property and violent index crimes rates.[2] As can be seen in all three crime curves, there was a general linear increase in the crime rates between 1960 and 1980. In 1960, the total index crime rate was 1,441 crimes per 100,000 California residents, but by 1980 this rate had increased all the way to 3,922 index crimes per 100,000 persons. The index crime rate in 1980 represented a 170% increase in crime compared with this rate in 1970. After 1980, however, the property and violent index crime rates show different trends. For property crimes, the crime rate decreased from 1981 through 1985, where it then leveled off for the next six years. Around 1991, the property crime rate began to decrease once again, a pattern that continued essentially until 2000, when this time series ends. The drop in the property crime rate from 1991 to 2000 is approximately 50 percent.

For violent offenses, however, the crime rate continued to climb from 1980 through 1992, where it peaked for the time series presented here. Thus, there was a general linear upward trend in the violent crime rate from the early 1960s through the

[2] The data displayed in Figure 6.1 through 6.3 were obtained from various data tables in *Crime and Delinquency in California, 1998*, *Crime and Delinquency in California, 1999*, and *Crime and Delinquency in California, 2000* (California Department of Justice 1999, 2000, 2001).

Figure 6.1. California Index Crime Rates from 1960 through 2000, by Crime Type

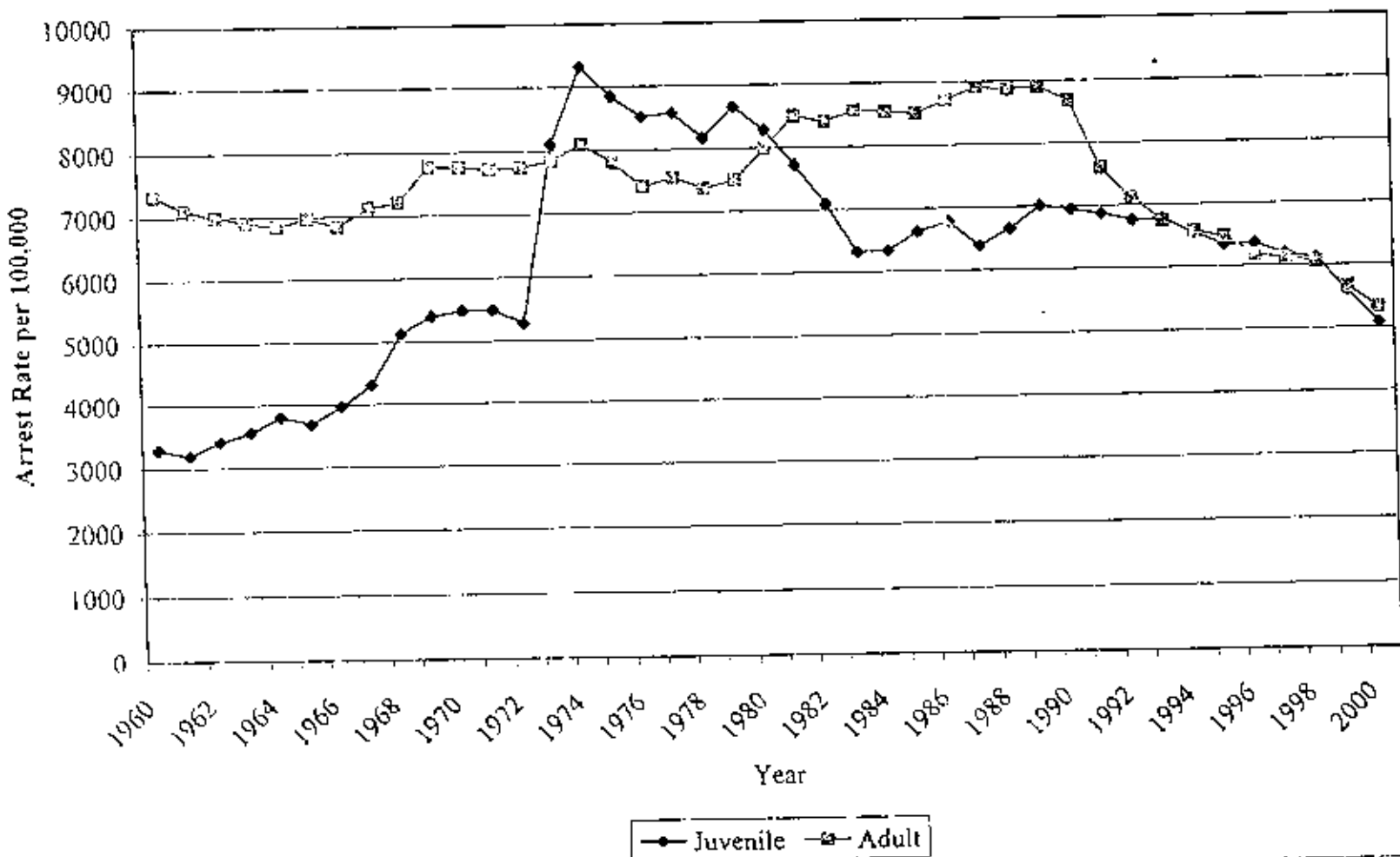


early 1990s. The upward trend in the violent crime rate between 1984 and the early 1990s enabled the total crime index trend to continue to increase slightly during this time period. Yet, between 1992 and 1999, California's violent crime rate decreased by almost one-half (a reduction of .44) from 1,103 violent crimes per 100,000 residents to 610 crimes per 100,000 persons. This rapidly decreasing crime trend during this time period was contrary to the expectations of several scholars who had predicted that waves of violent "superpredators" or "monsters" would be hitting and roaming the streets from the mid 1990s onward leading to a new crime wave (for more on this point, see Zimring 1998).

Consider next the reported arrest rates of both adults and juveniles in California (1960-2000) presented in Figure 6.2.[3] Compared to the juvenile arrest rate, the adult arrest rate from 1960 through the late 1980s displayed greater stability. The adult arrest rate slowly increased from 6843 per 100,000 residents in 1966 to 8900 per 100,000 residents in 1989, where it then began to descend until it reached 5323 per 100,000 persons in 2000. In fact, between 1989 and 2000, the total adult arrest rate decreased by over 40%. For juveniles, however, the arrest rate shows a somewhat different pattern. There was a general increase in the juvenile arrest rate between the early 1960s and the mid 1980s (especially around 1972-75). In fact, the arrest rate increased from 3300 per 100,000 residents (in 1960) to 9300 per 100,000 persons in 1975; this increase in the rates amounts to a 180% increase in the juvenile arrest rate. Between 1975 and 2000, the juvenile arrest rate would then decrease by nearly one-half (.46).

[3] For comparative purposes, note that the earliest year of arrest in the 1981-82 sample was 1965. The interquartile range (IQR) of the year of first arrest for the 1981-82 sample was {1973, 1978}. For the 1986-87 sample, the earliest year of arrest was 1970, and the IQR was {1979, 1982}. For the 1991-92 sample, the comparable numbers were 1974 and {1984, 1987}

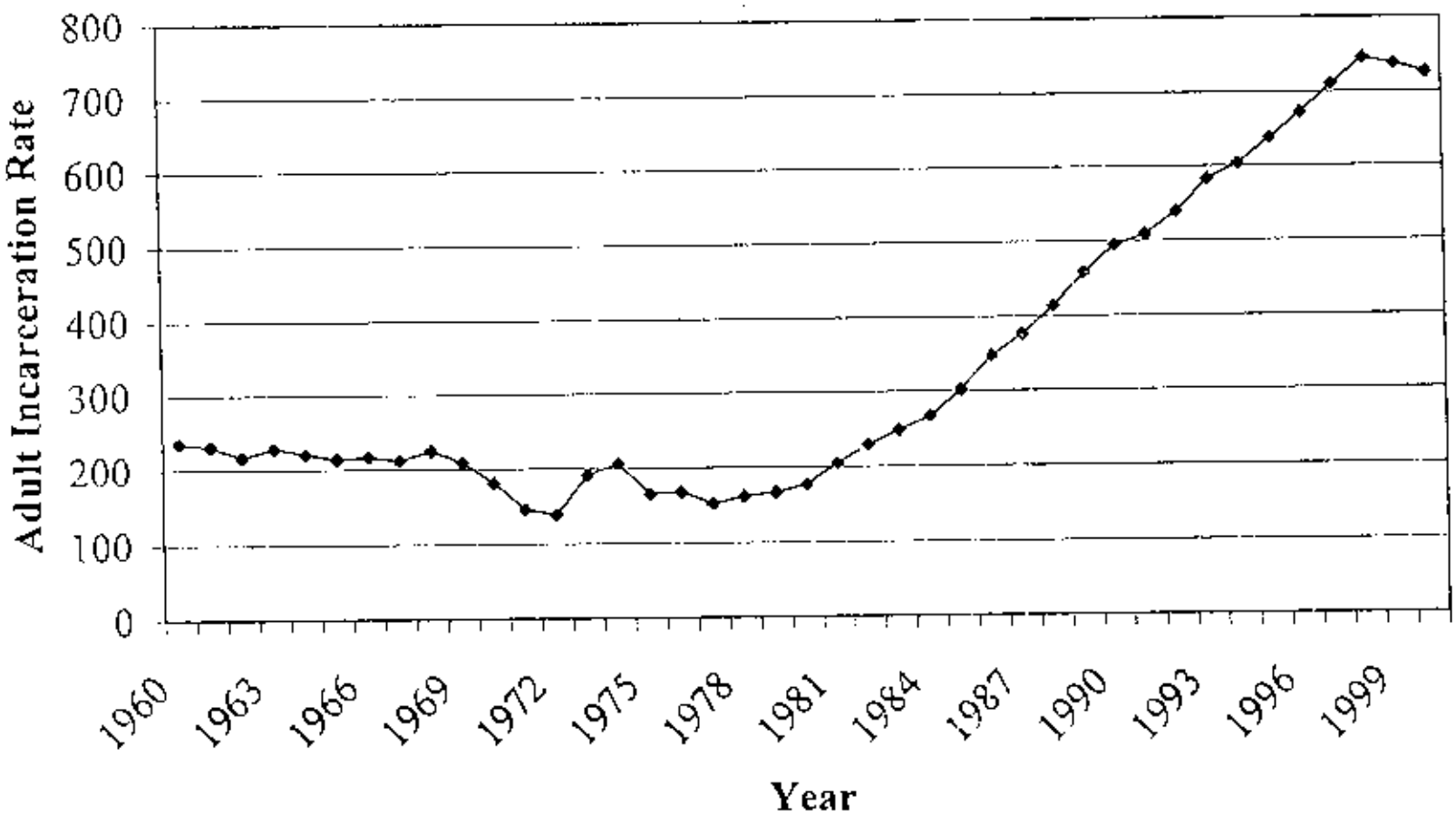
Figure 6.2 Juvenile and Adult Arrest Rates from 1960 through 2000



Between the 1960s and the end of the 20th century, both California and the United States itself went on what has been called an "imprisonment binge" (Irwin and Austin, 1997). During the decade of 1980-90, for example, the number of persons incarcerated in California's prisons and jails quadrupled, increasing the state's inmate population by over 100,000 prisoners (Zimring and Hawkins, 1994, 1997). Never before had "a prison system grown so much in so short a time" (Zimring and Hawkins, 1994: 83). Figure 6.3 contains a graph of California's adult incarceration rate (of CDC prisoners) between 1960 and 2000. As is clearly shown in this figure, there appears to have been a change in California's imprisonment rate around 1981 with respect to its increasing use of imprisonment in the CDC as a major form of punishment. Between 1981 and 2000, the imprisonment rate in California increased by 250%, with over 130,000 inmates having been added to the state's "prison industrial complex" known as the CDC.

The descriptions of the arrest and incarceration rates presented in the aforementioned figures are important because they help to convey the historical context in which a "get tough on crime" atmosphere was to impact the lives of wards who were to come under the supervision of the California Youth Authority. The members of each of the samples employed in our study may have been released at different points in time, however, each was impacted by a "get tough on crime" atmosphere that existed in the state and was directed at both adult and youthful offenders. In the next section, we show that not only were the members of our three samples deeply involved in serious crime, but that nearly all of these wards had been apprehended for engaging in some form of violent behavior at some point in their lives. In fact, the vast majority of these offenders had an arrest record for at least one serious violent offense.

Figure 6.3. California's Adult Incarceration Rate (CDC Prisoners) from 1960 through 2000



CRIMINAL ARREST HISTORIES

Up to this point, we have merely verbally indicated that the wards of the CYA comprise a rather select group of very serious offenders in the state. In this section, we present empirical evidence to buttress this characterization. In the first part of this section, we describe the age of onset for first criminal arrest for members of our samples. The second part of this section contains a description of the overall offending patterns of the three samples, including arrest counts, and the means, and prevalence rates for specific crime types.

Age at First Arrest

Age of onset is a critically important variable in the discipline of criminology because not only is it one of the most significant predictors of the frequency and seriousness of subsequent offending, but it is also a highly accurate predictor of the odds of becoming a persistent offender who continues offending into adulthood (see e.g., Loeber 1982; Farrington et al. 1990; Elliott 1994; Tolan and Thomas 1995; Tracy and Kempf-Leonard 1996; LeBlanc and Loeber 1998; Loeber et al 1998; Ayers et al. 1999; Lahey et al. 1999). Individuals with the earliest ages of onset (whether measured through self-report or official data) tend to have the longest criminal careers and to be involved in the most serious and violent offenses. They also commit criminal acts at higher rates than do others. Indeed age of onset is extremely important to the issues addressed in this study. Population heterogeneity theorists such as Gottfredson and Hirschi argue that an early age of onset is but a proxy indicator of higher criminal propensity, and that is why individuals with earlier ages of onset have criminals careers that tend to be longer, more

frequent, and more serious in nature (Nagin and Farrington 1992b; Dean et al. 1996). State dependence theorists, on the other hand, argue that an early age of onset could *cause* an individual to have a longer, more frequent, and more serious criminal career as a result of the negative (state dependence) effects generated by the early arrest (Nagin and Farrington 1992b). To the dual taxonomy theorists such as Moffitt and Patterson, age of onset is a marker variable that indicates whether the individual is in the life-course-persistent ("early onset") group or the adolescent-limited ("late onset") offender groups (Nagin and Farrington 1992b; Dean et al. 1996; Thornberry and Krohn 2001).

Table 6.1 contains the means and the 25th, 50th, and 75th percentiles of the distributions of the ages at first arrest for those in each sample and for all three samples combined. For comparative purposes, we also present the results on the basis of the court of commitment (juvenile court commitment and adult court commitments). For all 4866 cases in the three samples, the average age at first arrest was 13.89, and 75% of the wards in each of the samples had been arrested prior to age 16. The median age at first arrest (13.81) was very similar to the mean age, although the distribution is (not surprisingly) slightly skewed-right (which is diagnosed by the fact that the median is larger than the mean). In fact, a kernel density graph for the age at first arrest variable indicates that the distribution is fairly normally and symmetrically distributed around the mean, although it is slightly skewed-right.

More importantly though, there was not a large substantive difference in the mean (or median) age at first arrest among the three samples, although the 1981-82 sample cases were slightly older on average (14.2) than either the 1986-87 or 1992-92 samples (both averaged about 13.6). It is also very clear from these results that a sizeable

Table 6.1. Mean and Percentiles of Age at First Criminal Arrest Distribution, by Sample

	N	Mean	Percentiles 25th	50th	75th
All Wards					
1981-82 Sample	1989	14.22	12.55	14.24	16.01
1986-87 Sample	1443	13.68	12.31	13.71	15.11
1991-92 Sample	1434	13.64	12.49	13.81	15.07
All 3 Samples	4866	13.89	12.45	13.93	15.40
Juvenile Court Commitments Only					
1981-82 Sample	1068	12.98	11.70	13.29	14.51
1986-87 Sample	1089	13.21	12.01	13.39	14.55
1991-92 Sample	1308	13.50	12.42	13.74	14.92
All 3 Samples	3465	13.25	12.07	13.48	14.69
Adult Criminal Court Commitments Only					
1981-82 Sample	921	15.67	13.91	15.84	18.07
1986-87 Sample	354	15.14	13.67	15.09	16.93
1991-92 Sample	126	15.05	13.61	15.20	16.77
All 3 Samples	1401	15.48	13.85	15.58	17.54

percentage of each of the samples has a very precocious age at first arrest—roughly 25% of the samples have been arrested before they turned 13 years old. Of those who were arrested before age 13, about 50% were arrested before they turned 11 years old. Again, seventy-five percent of the cases in these three samples had been arrested prior to the age of 16.

When the samples were further disaggregated on the basis of the court of commitment, we found that the average age at first arrest for the juvenile court wards (in all three samples) was 13.25, whereas it was 15.5 for the adult court commitments. More importantly, after disaggregating the samples by court of commitment, the small differences that existed between the samples virtually disappeared. The reason the 1981-82 sample had a slightly higher mean age at first arrest was purely a consequence of the fact that this sample had a higher percentage of adult court commitments, who on-average had older ages at first arrest than did the juvenile court wards. Nonetheless, the main point indicated in Table 6.1 is that on average, the age at first arrest occurred fairly early in the lives of the members of each of our samples and also that there was some variation around this average age even in this select group of offenders.

Appendix D contains a graph depicting the cumulative ages of first arrest for each of the samples (Panel A in Figure D.1) and for the juvenile court commitments in each sample (Panel B in Figure D.1).

Criminal Arrest Patterns

Attention is now turned to the extent and nature of criminal activity in each of the samples. Table 6.2 contains information regarding the sum number of arrest charges, the mean number of arrest charges per case, and the participation rate for each offense type for each of the samples. The offenses have been broadly categorized into 6 main categories (which are in bold type in the Table 6.1): (1) Total; (2) Total Serious; (3) Violent; (4) Property; (5) Drug; and (6) Residual.[4] Attention here focuses on the results concerning the 6 main categories, but results are also presented in the table for 25 offense categories subordinate to the 6 main categories listed above. In total, the 4,866 cases in the three samples amassed 99,830 arrest charges as of June 30, 2000. Because the total time "at risk" varies greatly between the samples, we keep comparisons between the samples to a minimum.

The 1981-82 Sample

The 1,989 members of the 1981-82 sample accumulated 45,312 arrest charges. This averages out to just under 23 arrest charges per ward. Of the 23 average arrest charges, just over 50% of them (12.51) were for serious crimes such as homicide, robbery, aggravated assault, grand theft, auto theft, burglary, and drug trafficking. Disaggregating the arrests into violent, property, drug, and residual offenses, we see that these offenders are quite versatile in their criminal behavior, engaging in a wide array of offenses. On average, each of these wards had 5.3 arrest charges for violent offenses,

[4] Appendix A contains a description of the specific offenses used in generating these variables

Table 6.2. Sums, Means, and Participation Rates of Arrest Charges, by Sample

Charge Type	1981-82 Sample (N=1989)			1986-87 Sample (N=1443)			1991-92 Sample (N=1434)		
	Sum # of Charges	Mean # of Charges	Participation Rate	Sum # of Charges	Mean # of Charges	Participation Rate	Sum # of Charges	Mean # of Charges	Participation Rate
Total	45312	22.78	100.00	30776	21.33	100.00	23742	16.56	100.00
Total Serious	24885	12.51	99.85	16290	11.29	99.79	12321	8.59	99.93
Violent	10571	5.31	90.45	7601	5.27	91.61	6558	4.57	93.58
Serious Violent	6198	3.12	83.01	4313	2.99	81.57	3544	2.47	81.73
Homicide	217	0.11	9.15	158	0.11	9.42	148	0.10	9.07
Forcible Rape	356	0.18	11.97	222	0.15	10.53	161	0.11	8.37
Robbery	2869	1.44	60.99	1905	1.32	58.35	1458	1.02	52.79
Aggravated Assault	2071	1.04	51.58	1503	1.04	54.12	1359	0.95	55.44
Kidnap/Extortion	213	0.11	9.05	145	0.10	8.80	94	0.07	5.86
Child Molestation	132	0.07	4.88	114	0.08	6.31	68	0.05	4.18
Sodomy/Forced Oral Cop.	115	0.06	4.42	75	0.05	4.37	67	0.05	3.56
Weapon Discharge	225	0.11	9.35	191	0.13	11.43	189	0.13	11.09
Simple Assault	4373	2.20	64.91	3288	2.28	71.52	3014	2.10	69.46
Property	17337	8.72	95.48	10415	7.22	95.01	7242	5.05	90.31
Serious Property	14175	7.13	94.02	8297	5.75	92.52	6011	4.19	87.59
Burglary	6922	3.48	84.82	3722	2.58	79.63	2059	1.44	62.62
Auto Theft	3169	1.59	58.97	1977	1.37	55.79	2103	1.47	62.97
Major Theft	3938	1.98	67.87	2524	1.75	67.98	1773	1.24	57.53
Arson	146	0.07	6.33	74	0.05	4.44	76	0.05	4.18
Petty Theft	3162	1.59	61.49	2118	1.47	58.84	1231	0.86	44.21
Drug	9362	4.71	77.33	7115	4.93	80.39	4221	2.94	72.45
Serious Drug	4512	2.27	63.15	3680	2.55	68.68	2766	1.93	62.34
Sales/Trafficking	739	0.37	20.71	660	0.46	24.88	571	0.40	24.13
Posess./Poss. For Sale	3773	1.90	60.78	3020	2.09	65.49	2195	1.53	59.00
Other Drug Offenses	4850	2.44	55.66	3435	2.38	55.72	1455	1.01	38.01
Other Residual Offenses	8042	4.04	86.22	5645	3.91	90.02	5721	3.99	90.66
Felony Weapon Poss.	1379	0.69	37.91	989	0.69	38.46	836	0.58	37.80
Other Sex Offenses	337	0.17	10.06	193	0.13	8.18	158	0.11	7.88
Other Felony Offenses	540	0.27	19.26	410	0.28	21.83	522	0.36	26.71
Escape From Secure Fac	785	0.39	24.38	551	0.38	25.02	512	0.36	22.80
Other Misc Misd.	5001	2.51	76.77	3502	2.43	79.90	3693	2.58	81.31

Notes: Participation rate indicates the percentage of the cases in the samples with at least 1 arrest charge of that type.

"Serious Violent" includes all violent charges excluding simple assault charges. "Serious Property" includes all property charges except for the "Petty Theft" category. "Serious Drug" includes all drug charges except the "Other Drug Offenses" category. "Total Serious" includes the following: All "Serious Violent" charges, "Serious Property" charges, and "Serious Drug" charges.

8.72 arrest charges for property crimes, 4.7 charges for drug offenses, and about 4 charges for residual offenses.

Looking at a few of the sub-categories, we see that our sample members averaged about 3 serious violent offenses (which includes all of the violent offenses except simple assault), and they were arrested for 2 simple assault charges on average. Most of the property charges for this sample were also serious in nature. For example, they were arrested, on average, for 3.5 burglary charges. They were also arrested on average for 2.5 "miscellaneous" arrest charges (e.g., loitering, disturbing the peace, mail tampering, vandalism, tampering with an auto, gambling, possession of burglary tools, failure to appear in court, possession of false identification, and false information to a police officer).

The participation rates of the various crime types also speak to the wide variety of criminal activities in which the majority of these cases were involved at some point in their lives. Consider that over 90% of the sample members were arrested for at least 1 violent offense, and over 83% had at least one serious violent offense such as forcible rape, robbery, or aggravated assault. The most prevalent violent offense types were robbery (61%), aggravated assault (52%), and simple assault (65%). Almost every case in the sample was arrested for at least one serious property offense (94%), and over 85% of the cases had at least one arrest charge for burglary. Sixty percent of the cases had been arrested for auto theft. Three-quarters of the sample had at least one drug arrest charge in their arrest histories, and about 60% had been arrested at some point for either possessing illegal narcotics/marijuana or for possession these illegal commodities with the intent to sell; about 20% had been arrested for a drug trafficking charge. Finally,

about 86% of the 1981-82 sample had at least one arrest charge for the residual offense category.

The 1986-87 Sample

The cases in this sample (N=1,443) accrued 30,776 arrest charges by the end of the follow-up period, with an average of about 21 arrest charges per case. Similar to the 1981-82 sample, about half of their arrest charges were for serious offenses. These offenders were arrested on average for 5.27 violent offenses, 7.22 property offenses, about 5 drug offenses, and about 4 residual offenses. Looking at some of the specific offense categories in Table 6.2, we find that these individuals were arrested on average for 3 serious violent offenses, and just under 2.3 simple assault charges. With respect to their property offending, they were arrested on average for 5.75 serious property offenses, including an average of 2.6 burglary charges and 1.75 major theft charges (e.g., grand theft, forgery, possession of stolen property). This sample was also heavily involved in drug offenses, and indeed, even accrued more drug arrests (on average) than did the 1981-82 sample. On average, the1986-87 sample was arrested for two possession/possession for drug sale charges and 2.4 "other" drug offenses (e.g., under the influence of drugs, possession of drug paraphernalia, possession of hypodermic needle/syringe).

Again, an examination of the participation rates of the specific crime types confirms the extensive nature of the criminal activity of the 1986-87 sample. Ninety two percent of the sample had been arrested at least once for a violent offense, and about 82% of the sample was arrested for a serious violent offense. Roughly sixty percent was

arrested for a robbery charge, 54% for aggravated assault, and 72% for a simple assault offense. Like the previous sample, the members of the 1986-87 sample were also extensively involved in both property and drug crimes as well. Ninety-five percent of the sample was arrested for at least one property offense, 80% had a burglary charge, and similarly, 80% of the sample had been arrested for at least one drug charge; with seventy percent having been arrested for at least one serious drug charge. This sample was also heavily involved in the residual offense category as well, with 90% of the having at least one charge for offenses under this heading, with nearly 40% having been arrested on a felony weapons charge (e.g., possession of a machine gun, possession of a "sawed-off shotgun").

The 1991-92 Sample

Finally, Table 6.2 indicates that the 1991-92 sample was also heavily involved in criminal activities. The 1,434 members of this sample were arrested for 23,742 criminal charges. The "average" case in this sample was arrested for 16.56 charges, and once again, we see that over one-half of their arrest charges were for serious offenses. On average, they were arrested for 4.6 violent offenses, 5.1 property offenses, 2.9 drug offenses, and 4 residual offenses. Some of the arrest charge totals for specific offenses include an average of 2.5 serious violent charges, 2.1 simple assault charges, 4.2 serious property charges, 1.9 serious drug charges, and 2.6 "miscellaneous" charges.

The participation rates also indicate that the members of this 1991-92 sample were involved in a wide assortment of offense types. Indeed, 94% had at least one violent offense in their arrest history, 82% had been arrested for a serious violent offense-

- 53% for robbery, 55% for aggravated assault, and 70% had an arrest for simple assault. Again, we find that 90% of this sample had been arrested for a property offense, and that nearly an equal proportion of the cases (88%) also had at least one serious property arrest charge in their arrest history. Over 70% of the sample had been arrested on a drug charge, and 60% for had an arrest charge for either possession of drugs or possession of drugs with the intent to sell. Finally, about 90% of the 1991-92 sample had been arrested for a residual offense, including 38% for a felony weapon possession charge and 23% for an escape/attempted escape from a secure custody facility (e.g., jail, juvenile hall, CYA, CDC). Many of these escapes were rather dramatic. For example, one case was on his way to the county jail having already been arrested for burglary. While en route to the jail, the ward dove head first through the rear passenger's side window *while the window was rolled-up and the police car was moving* at about 25 miles per hour. This case was arrested a couple months later, at which point he was charged with misdemeanor theft of county property (the handcuffs) and escape. Another case involved a ward who had been arrested by his parole agent on a burglary warrant charge. While the parole agent was talking to the ward's parents, the ward jumped through a second story window *while handcuffed* and then escaped through a field in the backyard.

In sum, the results in Table 6.2 speak to the great quantity of criminal offending among these 4,866 individuals (who amassed almost 100,000 arrest charges in total or about 20 charges each), but these data also speak to the seriousness of their offending patterns. Almost all of the sample members had at least one serious offense in their arrest histories, and about half of all of the total arrest charges were for serious criminal

offenses.[5] Furthermore, and perhaps more importantly, over 80% of the members of each sample had been arrested for at least one serious violent offense. Indeed nearly 10% of each sample had been arrested at least once for homicide, and over half of the members in each sample had been arrested at least once for both aggravated assault (e.g., assault with a deadly weapon, aggravated assault on a peace officer) and robbery.

To conclude this section, we make a comparison to the Glueck's data utilized by Sampson and Laub (1993) to study crime among high-risk offenders. The Glueck's sample has been characterized as a group of serious and persistent offenders (Sampson and Laub 1993; Sampson 2000) and according to these authors (who also coded the three most serious charges per arrests), the 480 male juvenile delinquents had been arrested for 6300 criminal charges by age 32. This resulted in an average arrest charge total of 13.13 charges through age 32. For the two samples in our study that had a significant proportion of the sample at risk through age 32 (the 1981-82 and 1986-87 samples), we calculated the arrest charge totals through age 32 for the proportion of each sample that was at risk to this age. For the data used here, the comparable arrest charge averages were 21 in the 1981-82 sample and 20.91 in the 1986-87 sample. Thus, the offenders in our study average around 7 more arrest charges by age 32 in comparison to the juvenile delinquents in the Glueck's data used by Sampson and Laub.

[5] The handful of cases not arrested for a serious offenses were chronically involved in either petty theft and/or simple assault).

Post-Release Recidivism Rates

In this section we discuss the post-release recidivism rates for each of the parole release samples. Here, we use arrest for a new criminal charge as our measure of recidivism, whereas the CYA generally reports parole removal rates (which include parole revocations, CYA recommitments for law violators, and discharges due to jail and prison commitments for law violations) in the 24 months since release as its measure of recidivism. We note that on May 5, 1998, the Director of the CYA at that time (Francisco Alarcon) released the following press release:

> The recidivism rate for offenders released from the California Youth Authority (CYA) has reached its lowest level in nearly two decades, according to a parole performance report released by CYA Director Francisco J. Alarcon.
>
> The report, which is based on findings from 3,212 offenders released to parole in 1995, shows that less than half (49.6 percent) violated their parole within 24 months - down from a 55.1 percent recidivism rate for those offenders released in 1994. In addition, the report reflects the largest single year drop from 55.1 percent last year to 49.6 percent for the 1995 releases. The parole failure rate for females dropped to 24.8 percent, its lowest since 1979, and the rate for males declined to 50.6 percent, down from a high of 61.7 percent in 1986.
>
> "These numbers reflect an improved success rate for CYA parolees, and that translates into a positive impact on public safety and hundreds of fewer victims of crime," said Alarcon. "We incarcerate the most serious and violent young offenders in California, and the fact that we're effectively reaching over half of them speaks to the quality of our staff and our institutional programs"
> (http://www.cya.ca.gov/Library/news/recidvsm.html, 9/17/02).

These statistics, however, must be interpreted with caution because they rely on the CYA's basic measure of recidivism—the 24-month parole outcome. In our opinion, this 24-month parole performance measure is a flawed measure of recidivism. It includes

data only for the period in which wards are under the supervision of CYA parole agents. Over the past two decades, an increasing proportion of CYA wards have been released with little or no jurisdiction time available for parole supervision (i.e., their period of parole supervision does not cover the full 24 months). Therefore, we believe that actual criminal arrests are a better indicator of the post-release performance of these wards because arrest data are not dependent on the length of parole supervision, and thus are a better indicator of post-release performance. Further, as Maltz (1984) argues, failure on parole can arise for many reasons, some of which are not even violations of the criminal law (e.g., failure to obtain a job). Thus, we believe that parole failure is not a very precise measure of subsequent criminal activity, since one's parole can be revoked for a variety of non-criminal reasons.

Using subsequent arrests for a new criminal charge as our measure of recidivism, Table 6.3 contains the estimated and actual probability of surviving or remaining "arrest free" for various periods of post-release time. The estimated survival probability columns represent the estimated probability of "surviving arrest free" up to at least the specified time period in each row (representing each year after release). This statistic was obtained using the Kaplan-Meier product-limit formula defined as

$$\hat{S}(t) = \prod_{j:t_j \le t} \left[1 - \frac{d_j}{n_j} \right],$$

where n_j represents the number of individuals at risk at time t_j (i.e. those not arrested yet and still under observation) and d_j represents the number of individuals at risk who were arrested during time t_j (Allison 1995). The quantity in the brackets represents the conditional probability of remaining arrest free during a given interval of time given that

Table 6.3. Estimated Survival Probability of Remaining "Arrest Free" of a New Criminal Charge as a Function of Years Since Release

	Estimated Survival Probability in Each Sample		
Years Since Release	1981-82	1986-87	1991-92
1	0.502	0.458	0.482
2	0.342	0.299	0.306
3	0.255	0.217	0.225
4	0.210	0.162	0.171
5	0.180	0.128	0.136
6	0.160	0.117	0.114
7	0.147	0.106	0.107
8	0.135	0.095	0.098
9	0.128	0.089	--
10	0.122	0.084	--
11	0.119	0.081	--
12	0.113	0.080	--
13	0.111	0.077	--
14	0.107	--	--
15	0.105	--	--
16	0.103	--	--
17	0.096	--	--
18	0.094	--	--
Estimated Median Survival Days	368	309	344
Observed Survival Rate	0.102	0.086	0.111

Note: The observed survival rate reflects the actual probability of remaining arrest free (i.e., no new criminal arrest charge) through June 30, 2000. It does not account for the fact that some cases were not at risk of arrest over the entire follow-up period (i.e., some cases died before the end of the follow-up period). The estimated survival probability presented above properly accounts for the effects of this "censoring" of at risk time.

the individual has remained arrest free up to that point in time (Allison 1995). The product-limit estimate represents the product of these conditional survival probabilities up to a specified period of time. This method of calculating the survival probability properly accounts for the effects of the "right-censoring" in this data that results from the fact that not all individuals were at risk during the entire follow-up period because some of the subjects had died (as described in a subsequent section below) and therefore were not at risk of being arrested. Due to the fact that some individuals were arrested at times later than the death (or censoring) times of other individuals, the observed survival probability at the end of the follow-up period (June 30, 2000) is biased upwards (or overestimated).

In Table 6.3, we find that there was a considerably high rate of arrest during the first year of parole release for all three samples. In fact, roughly 50% or more of each sample was arrested for a new criminal arrest charge within the first 365 days after release on parole. More specifically, the estimated probability of surviving arrest free past the first year of release was 0.502 in the 1981-82 sample, 0.458 in the 1986-87 sample, and 0.482 in the 1991-92 sample. The estimated median number of days until arrest ranged from 309 in the 1986-87 sample to 368 in the 1981-82 sample, with the 1991-92 sample having an estimated median time until arrest of 344 days. By the end of the third year, roughly 75% or more of each sample had been arrested for a new criminal charge. By the end of the eighth year of follow-up period (the last point in time at which all three samples were measured), the estimated survival probabilities were each at around 0.10 (about 10% had not been arrested yet), indicating approximately a 90% failure rate. The 1981-82 sample had the highest estimated probability of surviving arrest

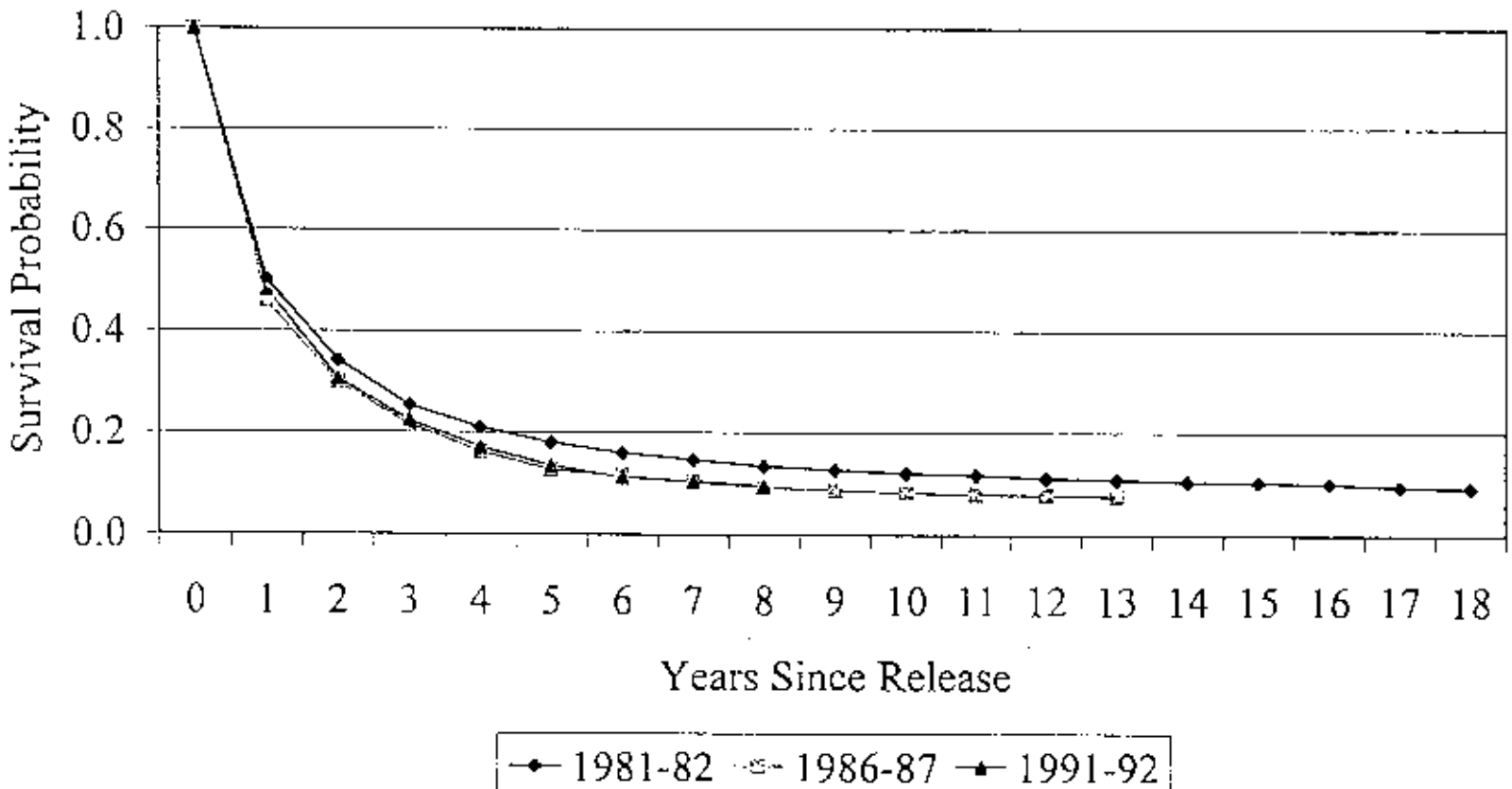
free for at least eight years (13.5%), and the other two release samples were just under 0.10 (0.095 for the 1986-87 sample; 0.098 in the 1991-92 sample). Indeed, by the end of the follow-up period (June 30, 2000), the estimated probability of surviving arrest free was under 0.10 in all three samples.

In Figure 6.4, we depict a graphic portrayal of the survival probabilities presented in Table 6.3. There are several interesting, illustrative points concerning the recidivism rates of the three samples that can be gleaned from the survival curves presented in this figure. First, the survival curves presented in Figure 6.4 clearly indicate that the CYA wards in all three samples were at an exceptionally high-risk of arrest for a new criminal charge during the first year of release (i.e., the drop in the curves is exceptionally steep in the first year). Second, it is also clear that there was a significant decline in the rate of a first arrest (for a new criminal offense) after the third year of release given that an individual was able to remain arrest free for the first three years (i.e., the curves start to level off, although they are still dropping in each year). Third, the curves presented in Figure 6.4 also make it blatantly clear that relatively few individuals in any of the three release samples were able to remain arrest free during the follow-up period. In fact, we believe that these results point to the inadequacy of the CYA's generally employed recidivism measure (parole removal within 24 months of release), and that the more adequate measure of recidivism rates of CYA ward's employed here indicate that the recidivism rates of individuals released from the CYA are considerably higher than those presented by the CYA in its publications.[6]

---

[6] This point is supported to an even greater degree by the fact that more than 50% of the cases in each of these three samples end up incarcerated in the CDC for a new conviction for a serious criminal charge that warranted incarceration in the state penal system.

Figure 6.4. Estimated Survival Probabilities of Remaining Arrest Free as a Function of Years Since Release



Finally, with respect to comparing the survival curves of the three samples, it is evident from these curves (which connect the survival probabilities at the various time points) presented in Figure 6.4 that the survival curves of the 1986-87 and 1991-92 samples were nearly identical, while the 1981-82 sample had a curve that was marginally elevated above the other two samples (indicating a *marginally* better chance of remaining arrest free). However, we show in Chapter 7 that there is a "class" of offenders in the 1981-82 sample that was unique to this sample and was composed of a group of males who: (1) had a later age at first arrest, (2) had a very low rate of arrest over time, and (3) were largely sentenced from the adult criminal courts. This class or group of offenders was removed from the CYA population as a result of the legislative changes described in the Chapter 4. Removing this group of offenders from the sample used to calculate the survival probabilities presented in Table 6.3 and Figure 6.4 resulted in nearly identical survival curves for each of the three datasets. In fact, after removing this class of offenders from the 1981-82 sample, the survival curves across time for all three samples collapsed on one another to the point that they could have been represented by a single survival curve. In other words, the survival probabilities were essentially identical in all three samples—the marginal survival curves indicate that there were no significant differences in the recidivism rates between these three samples, especially after removing the one class of offenders unique to the 1981-82 sample (and who were removed from the CYA population purely as a result of legislative changes).

BACKGROUND CHARACTERISTICS VARIABLES

The data displayed in Table 6.4 depict the background characteristics of the wards in our study at the time they were incarcerated in the CYA for the sample stay. All of the variables in Table 6.4 are binary or dummy variables, and thus the means displayed in the table represent the proportion of the sample coded as 1 on the variable (which was used to indicate the presence of the characteristic). Again attention here focuses on providing a description of the cases, and not explaining the presence and/or significance of the differences between the samples. As a general comparison of the differences between the samples, we note that the overall picture of the results regarding our data (excluding the ethnicity variables) indicates that, in terms of the background characteristics, the 1986-87 sample came into the CYA with a "more troubling past" than did the 1981-82 sample. Similarly, the 1991-92 sample came into the CYA in a "more troubling past" than did the 1986-87 sample. For some of the "troubling characteristics," the greatest change occurred between the 1981-82 and 1986-87 samples, while for other such characteristics the greatest change was between the 1986-87 and 1991-92 samples.

The first set of results we report here pertains to the ethnicities of the wards that we have categorized here into 4 broad categories: White, Hispanic, African-American, and "Other" (which is primarily composed of individuals of Asian descent). In the 1981-82 sample, White and African-American wards each consist of about 35% of the sample, the Hispanics comprise 25%, and the remaining 2% were in the "Other" category. In the other two samples, the percentage of White wards decreased in each sample (28% of the 1986-87 sample and 20% of the 1991-92 sample). The Hispanic wards made-up 33% and 34% of the 1986-87 and 1991-92 samples, respectively. The percentage of cases that

Table 6.4. Means of Background Characteristic Variables, by Sample

Variable	Sample		
	1981-82 (N=1989)	1986-87 (N=1443)	1991-92 (N=1434)
Ethnicity			
White	0.36	0.28	0.20
Hispanic	0.25	0.33	0.34
African-American	0.36	0.37	0.41
Other	0.02	0.02	0.06
Family Background Characteristics			
Family Violence	0.17	0.18	0.23
Parental Alc./Drug Dependence	0.27	0.34	0.42
Parental Criminality	0.20	0.28	0.33
Sibling Criminality	0.39	0.38	0.43
Lack of Supervision/Neglect	0.35	0.43	0.65
Ineffective Control	0.60	0.77	0.86
Subject Characteristics			
Physical Abuse	0.15	0.20	0.22
Sexual Abuse	0.01	0.03	0.05
Drug Abuse	0.69	0.86	0.72
Gang Member/Association	0.29	0.51	0.74
Previous Violent Behavior	0.80	0.86	0.88
School Dropout	0.55	0.39	0.87

Note: All variables are binary (dummy) variables and the means represent the percentage of cases coded as 1 (which indicates the presence of the characteristic).

were African-American also increased over the two samples, constituting 37% (1986-87) and 41% (1991-92) of these samples respectively. The percentage of cases in the "Other" category was stable between 1981-82 and 1986-87 (2%), but increased to 6% in the 1991-92 sample.

The next set of data in Table 6.4 concerns what we have previously termed "troubling characteristics" concerning the ward's family environment. About 20% of each sample was raised in a family setting in which there were reported acts of family violence and (roughly speaking) around one-third of the wards in each sample had one or more parents with a drug or alcohol dependence problem (27% in 1981-82; 42% in 1991-92). There is also some evidence of relatively high rates of criminal activity of both parents (20-33%) and siblings (around 40% in all three samples), and a substantial number of the wards were raised in family settings characterized by either a lack of supervision or neglect of the ward (35%-65%). It was even more likely that the wards in the samples were reported to have been beyond the control of their parents/guardians (60% in 1981-82; 77% in 1986-87; 86% in 1991-92).

The final set of data presented in Table 6.4 relates to some of the other "troubling characteristics" in the social backgrounds of the subjects. Around 20% of the wards were reported to have been subject to physical abuse/extreme punishment when they were growing up, and about 5% or less of the cases in each sample had reportedly been sexually abused. The recorded drug abuse patterns also speak to a major problem or "troubling characteristic" afflicting a great majority of the wards in each sample. About 70-80% of the wards in these samples were reported to have had a history of drug abuse problems in their backgrounds, with some of these wards addicted to heroin,

cocaine/crack, and sniffing paint. Indeed, one case in the 1991-92 sample was already a heroin and methamphetamine addict at the precocious age of ten, after having been introduced to these drugs at age 9 by an uncle. The estimates of drug abuse patterns reported in Table 6.4 are right near the estimated percentage of the CYA population with a drug abuse problem (70-80%) based on a detailed questionnaire/interview with a sample of CYA wards (see, Haapanen and Ingram 2000; Wilson et al. 2001). Our data indicate that there was a precipitous increase in the percentage of each sample that was either reported to be a gang member or was a known "associate" of a gang, and this increase (from roughly 30% in 1981-82 to 75% in 1991-92) coincides with the documented proliferation of street gangs and the explosion in gang-related homicides occurring in the period between the early 1980s through the early 1990s (Klein 1995; Maxson 1999). Lastly, a considerable number of the wards in these samples had already dropped out of school prior to their incarceration for the sample stay period.[7] This was especially true for the 1991-92 sample—87% of the wards in this sample had already dropped out of school at the time of their incarceration.

VARIABLES RELATED TO THE SAMPLE STAY

A summary description of the variables related to the sample stay is presented in Table 6.5. As we have seen in Table 6.1, there was a dramatic difference between the samples in terms of the percentage of the cases committed from the juvenile court. This, of course, is a direct result of the legislation passed in the early 1980s (as described in

[7] Some of the cases had voluntarily dropped out or had just stopped attending without formally withdrawing. Other cases had been formally expelled from the school district in which they resided, often times due to an arrest for possession of a loaded firearm on school grounds.

Table 6.5. Means of Variables Related to Sample Stay At CYA, by Sample

Variable	Sample		
	1981-82 (N=1989)	1986-87 (N=1443)	1991-92 (N=1434)
Juvenile Court Commitment	0.54	0.75	0.91
Admission Status			
First Commitment	0.79	0.58	0.70
Parole Violator	0.11	0.36	0.26
Recommitment	0.09	0.06	0.04
Age at Admission	18.34	18.58	17.88
Age at Release	19.45	20.10	19.67
Length of Stay (in months)	13.43	18.28	21.48
Commitment Offense			
Violent	0.42	0.39	0.40
Property	0.53	0.48	0.39
Drug	0.02	0.08	0.16
Other	0.03	0.05	0.04
County (of Commitment) Region			
Los Angeles	0.40	0.45	0.41
Bay Area	0.20	0.19	0.14
Southern California (not LA)	0.19	0.12	0.15
Northern/Central California	0.22	0.24	0.30
Number of DDMS Infractions	1.09	1.92	2.01

Note: All variables are binary (dummy) variables except for age at admission, age at release, length of stay, and number of DDMS infractions. For the binary variables, the means represent the percentage of cases coded as 1 (which indicates the presence of the characteristic).

Chapter 5) that put constraints on the CYA as a sentencing alternative for young adults between the age of 18 and 21. The majority of each sample was in the CYA for a first commitment (60-80%), but the 1986-87 and 1991-92 samples had a larger number of parole violator admissions than did the 1981-82 sample. Recommitment wards were the least prevalent admission type (at around 5-10% of each sample). The average age at admission for the sample stay was about age 18, and on average the wards were between the ages of 19.5 and 20 years old when they were released from the CYA. There was a large increase in the average length of stay between the three samples (as noted in Chapter 4. The average offender in the 1991-92 sample was incarcerated about 21.5 months, whereas the average ward in the 19 81-82 sample was incarcerated for thirteen months. The majority of the wards were committed to the CYA for either a violent (40%) or property (40-50%) crime. There was an increase in the percentage of wards committed for a drug offense (2% in 1981-82 and 16% in 1991-92), which is arguably a consequence of the "War on Drugs" movement that began in the early 1980s.[3] Most of the wards in these samples came from either Los Angeles county (40%) or from Northern/Central California (22-30%) counties such as Sacramento, Fresno, and Kern. The remaining cases were either from counties in the Bay Area (14-20%) such as Alameda, San Francisco, and San Jose or from counties in Southern California (excluding Los Angeles) such as San Diego, Ventura, and San Bernardino (12-19%). Finally, the wards in these three samples were written up for an average of 1.1, 1.92, and 2.02 DDMS

[3] Some of the drug arrests that resulted in the commitment to the CYA involved small amounts of drugs such as a few individually packaged "rocks" of crack cocaine, while other times it involved large amounts of narcotics. One case was arrested for selling one and a half pounds of black tar heroin to undercover DEA agents, another was arrested with 55 "balloons" of heroin, while another case was arrested while in possession of $36,000 worth of crack cocaine, an automatic firearm, and $5,000 in cash.

infractions respectively during their sample stay at the CYA. The increasing infraction counts over time is probably the result of two characterizations of these samples already noted: the increasing average amount of time served mandated by YA Parole Board over time and the increase in gang membership among clients served by the YA over time. Those two characteristics tend to increase the chances for fights and assault among wards, and many of these fights and assaults reported among the 1991-92 sample were described as "gang related."

MORTALITY EXPERIENCES

Attention is now turned to the mortality experiences of the members of the three samples. Table 6.6 presents the frequencies of death events in each of the samples, and decomposes the major types of death into "High Risk" and "Other Types."[9] The "High Risk" death types included homicide, suicide, drug overdose, AIDS, and auto accident deaths.

Of the 4,866 individuals in the three samples, 329 (6.8%) were found to be deceased at some point after their release, with 152 (7.6%), 98 (6.8%), and 79 (5.5%) deaths occurring in the 1981-82, 1986-87, and 1991-92 samples, respectively. Even more importantly, about eighty percent or more of the deaths were found to be in the "High Risk" death types categories. That is, the majority of the deaths in the sample were not the result of "natural causes," but rather most were the result of high-risk activities.

[9] The major type of death was obtained from the ICD9 and ICD10 (1999 deaths) codes listed in the DSMF files. The ICD9 codes for firearm homicides included the following codes: E965.0 through E965.4 and E970. The ICD9 non-firearm homicides included the following codes: E960 through E964, E965.5 through E969, and E971 through E978. The homicide deaths in the 1999 DSMF file were identified using the ICD10 "group codes" 338 through 346, and group code 340 was used to determine if it is was a firearm homicide

Table 6.6. Frequency and Type of Death, by Sample

	Sample								
	1981-82 (N = 1989)			1986-87 (N = 1443)			1991-92 (N = 1434)		
Death Type	n	% of Deaths	% of Sample	n	% of Deaths	% of Sample	n	% of Deaths	% of Sample
All Death Types	152	100.0	7.6	98	100.0	6.8	79	100.0	5.5
"High-Risk" Deaths	117	77.0	5.9	83	84.7	5.8	69	87.3	4.8
Homicide	62	40.8	3.1	60	61.2	4.2	61	77.2	4.3
Firearm	51	33.6	2.6	51	52.0	3.5	58	73.4	4.0
Non-Firearm	8	5.3	0.4	8	8.2	0.6	2	2.5	0.1
Legal Intervention	3	2.0	0.2	1	1.0	0.1	1	1.3	0.1
Suicide	8	5.3	0.4	5	5.1	0.3	3	3.8	0.2
Drug Overdose	22	14.5	1.1	11	11.2	0.8	3	3.8	0.2
AIDS	13	8.6	0.7	4	4.1	0.3	0	0.0	0.0
Auto Accident	20	13.2	1.0	8	8.2	0.6	5	6.3	0.3
Other Types of Death	35	23.0	1.8	15	15.3	1.0	10	12.7	0.7

Descriptions of four of the deaths in the 1991-92 sample can help to describe the high-risk nature of some of these deaths.

> Case 1: Decedent was shot in the head during a shoot-out with an armed security guard during the robbery of a store. Decedent was shot after wounding the guard.
>
> Case 2: Decedent choked on his own vomit during a drug overdose.
>
> Case 3: Decedent was shot execution-style with the muzzle pressed against his forehead. Fellow gang member was arrested for the offense. Decedent was killed for testifying against the homicide perpetrator.
>
> Case 4: Decedent lost control of motorcycle at over 100 miles per hour.

Perhaps the most striking feature of Table 6.6 concerns the homicide event totals—of the 329 deaths, 183 of them (or 56%) were homicide victims! Furthermore, there was a considerable disparity across the samples in terms of the percentage of deaths that were attributable to homicide. Consider that the 1991-92 sample had ten less years of exposure and roughly 550 fewer wards than the 1981-82 sample, and yet the 1991-92 sample accrued almost the exact same count number of homicides as did the 1981-82 sample. In the 1981-82 sample, 40.8% (n = 62) of the deaths were attributed to homicide, whereas in the 1991-92 sample 77.2% (n = 61) of the deaths were reported as the result of homicide. The 1986-87 sample fell nearly in the middle of the other two samples—61% (n = 60) of the deaths in this sample were the result of homicide. Of the 183 homicide events in all three samples, 160 of them (or 87%) were the result of gunshot wounds.

The extent of homicide mortality in these three samples should not be taken lightly. To make a comparison concerning the heightened risk of mortality among the

samples, consider that the homicide rate peaked in California in 1993 at 12.9 homicide events for every 100,000 individuals. Thus, to obtain 183 homicide events in the general population, one would have needed a sample of about roughly 1.4 million people for that year. Granted the three samples analyzed herein were followed over time, but the fact that 183 homicide events were recorded in a sample of only 4866 individuals is in our view, an extreme homicide rate under any circumstances.[10]

ADULT INCARCERATION EXPERIENCES

Finally, Table 6.7 presents summary information regarding the entrances of sample cases into the California Department of Corrections (CDC). Again, the different amounts of exposure time between the samples should be kept in mind when interpreting the results in Table 6.7. As can be seen in the table, a sizeable percentage of each of the samples had been incarcerated in the CDC at least once as of June 30, 2000—65% of the 1981-82 sample, 69% of the 1986-87 sample, and 52% of the 91-92 sample. As of June 30, 2000, 13% of the 1981-82, 12% of the 1986-87, and 9% of the 1991-92 samples respectively had been sentenced as "Two Strikers" under California's "Three Strikes And You're Out" program. A larger percentage (3.3%) of the 1981-82 sample had been sentenced as a "third strike" case than either the 1986-87 sample (2.1%) or the 1991-92 sample (0.6%). This is not surprising since the selection mechanisms built-into the "Three Strikes" law in California put older individuals with convictions in their past at the greatest risk of being sentenced as a "Three Strike" case. Thus we would expect the

[10] For example, out of the 144,245 individuals who died between 1990 and 1999 in California that were born in the same years as the members of these three samples (i.e., between 1956 and 1978), only 15% of these individuals died as a result of homicide. In the three CYA samples, however, 55% of the cases that died between 1990 and 1999 died as a result of homicide.

Table 6.7. Summary of Variables Related to Adult Incarceration Experiences, by Sample

Variable	Sample		
	1981-82 (N=1989)	1986-87 (N=1443)	1991-92 (N=1434)
%With At Least 1 CDC Stay	65.1%	68.9%	52.3%
Maximum Strike Status:			
"Second Striker" %	12.7%	11.7%	9.1%
"Third Striker" %	3.3%	2.1%	0.6%
Among Those With at Least 1 CDC Stay:			
Avg. # of CDC Stays	2.5	2.1	1.6
Avg. Length of Stay (in months)	25.5	22.2	19.5
Median Length of Stay (in months)	16.3	14.5	13.8
Average Age at Admission	27.6	26.2	23.7
Average Age at First Admission	24.2	24.2	22.8
Avg. # of Total Years Incarcerated in CDC	5.4	3.9	2.5
Median # of Total Years Incarcerated in CDC	4.4	3.2	2.0

Note: Length of stay was calculated as of June 30, 2000 if they were still incarcerated at that point.

1981-82 sample to have the most individuals sentenced as a third strike case since they have had a longer time period over which to accrue such a legal status.

Limiting the sample to only those individuals who had at least one stay (or incarceration) in the CDC, we next generated some variables for these specific individuals. The average number of entrances among those with at least one CDC entrance was 2.5 (for the 1981-82 sample), 2.1 (for the1986-87 sample), and 1.6 (for the1991-92 sample). The average length of stay ranged from 20-25 months, while the median length of stay ranged from 14-16 months.[11] The average age of all admissions ranged from 23.7 (1991-92) to 27.6 (1981-82), and the average age at first admission was 24.2 in both the 1981-82 and 1986-87 samples, and 22.8 in the 1991-92 sample. The cumulative average number of years spent incarcerated in the CDC (as of June 30, 2000) was 5.4 years in the 1981-82 sample, 3.9 years in the 1986-87, and 2.5 years in the 1991-92 sample, with the median cumulative number of years incarcerated ranging from 2 to 4.4 years. Among those who were incarcerated at least once (and taking into account the differing amounts of exposure time in each sample), the median number of cumulative years incarcerated was just under one-quarter (about .24) of the total post-release period in each sample.

[11] The distribution of length of stay is right-skewed due to the effects of the "lifers" (those sentenced to life for murder or other extremely serious offenses) and those individuals incarcerated for long periods of time. Thus, we have presented both the median and the mean lengths of stay as of June 30, 2000. The median is a more accurate depiction of the typical sentence because of the skewed nature of the length of stay distribution

SUMMARY

The main point we hope to have conveyed in this chapter is that the wards in our samples constitute by any stretch of the imagination a group of very serious offenders. To put it bluntly, when it comes to accumulating a record for serious crime (for any comparable period of time), the members of the samples of the other studies we have reviewed here appear not to be "from the same league" as the members of our samples. Not surprisingly, many of the wards in our samples had led seriously troubled lives on a number of fronts and appear to have faced a myriad of major problems (besides the problems associated with their criminal activity). Drawing on a term from medical sociology, considering their crime, social problems (e.g., drug abuse, school dropout), and mortality risk simultaneously, the individuals in these three samples could be accurately characterized as having serious *co-morbidity* problems. In other words, while these individuals frequently engage in a wide variety of crimes throughout their life course, crime is not the only problem in the lives of these wards. Many of the individuals in these samples were raised in rather chaotic family environments, had serious academic difficulties and educational deficits, and many of them (especially in the 1991-92 sample) were raised in violent neighborhoods that were entrenched in gang activity. For many of the individuals in these samples, their stay at the CYA as a juvenile/young adult was a mere prelude to an accumulation of additional serious problems in adulthood, including future incarceration in the CDC for some and, unfortunately, a much higher than average risk of death and homicide compared to others in the general population (which usually was a result of being shot with a firearm).

Having now completed a summary description of the data, we next turn the focus of attention to the two main substantive chapters of this study. In the next chapter, we focus on the relationship between age and crime within latent classes of serious youthful offenders.

CHAPTER 7

AGE & CRIME AMONG LATENT CLASSES OF SERIOUS YOUTHFUL OFFENDERS

INTRODUCTION

In the first chapter of this study it was noted that one of the foremost controversies in contemporary theoretical criminology concerns how to simultaneously explain the sources of both continuity and change in criminal behavior over time. The "paradox of persistence" was described as a descriptive empirical finding that highlights the importance of the processes of both continuity and discontinuity in criminal offending patterns across the life course. In the first and second chapters, we detailed how the crux of the dispute between three different broad theoretical paradigms (i.e., population heterogeneity, state dependence, and the dual taxonomy approach) centers largely on predicted differences regarding the degree to which criminal propensity is stable/instable across the life course. The explanation as to why criminal propensity is either stable/instable (or a mixture of both) is of fundamental importance to the discipline of criminology because it has important theoretical implications for two of the robust or "brute facts" of criminology: (1) the relationship between age and crime; and (2) the association between past and subsequent criminal activity. We examine the topic of the association between past and subsequent criminal activity in Chapter 8, but first we investigate here the empirical relationship between age and crime within latent classes of offenders (that are homogenous with respect to their criminal activity across the life

course).[1] We presented four hypotheses in Chapter 3 that address the relationship between age and crime, that included investigating (1) whether there are multiple types of offenders, even on the high-end of the criminal propensity continuum; (2) whether there are specifically only two different offender types corresponding to the "adolescent-limited" and "life-course-persistent" groups; (3) whether there exists an "adolescent-limited" offender type in the high-risk samples used herein; and, finally, (4) whether there are stable between-group differences across the age range. Prior to presenting the empirical results of this chapter with respect to these issues, we first briefly address the age-crime relationship implied in the theories of Gottfredson and Hirschi (1990), Sampson and Laub (1993), and Moffitt (1993), which respectively represent the population heterogeneity, state dependence, and the dual taxonomy approaches to the explanation of criminal activity across the life course.

Recall that the population heterogeneity theory of Gottfredson and Hirschi (1990) posits that between-individual variation in criminal propensity is sufficient to explain not only the criminal career "parameters" of onset, prevalence, incidence, and duration/termination, but also the relationship between age and crime in general. According to Gottfredson and Hirschi (1990) all offenders decrease their offending over time. However, the between-individual differences in criminal propensity that exist at any one point in time (which become relatively fixed around the age of 8) continue to exist at all other points in time. The shape of the age-crime curve is hypothesized to be robust from person-to-person (i.e., the shape is invariant). Stated differently, individuals differ

[1] The results of this chapter are also critical for Chapter 8. We make use of the latent class results presented in this chapter in several of the analyses in Chapter 8.

with respect to their mean rates of offending, but the age parameters governing the shape of their offending trajectory are the same for every individual in the population. This argument that the parameters of the age-crime relationship are identical across persons is the core of Gottfredson and Hirschi's age invariance argument. All people, everywhere, and within any historical period, tend to commit less crime as they age. According to Gottfredson and Hirschi then, all individuals have their greatest involvement in criminal activity during the late adolescent years of life, and their offending rates decline uniformly thereafter. Again, the key implication of this argument is that *differences between individuals persist over time*, or stated in terms of the crime-criminality distinction, both the relative differences of criminal propensity (criminality) and relative group differences in criminal offending (crime) endure throughout life. The invariance argument and the stability of differences across the life course have profound implications for the study of crime. For example, if the relationship between age and crime is invariant (and between-group differences that exist at one point in the age-crime curve continue to exist at any other point in the age-crime curve), then only a single time point (cross-sectional data) will be sufficient to measure the criminality of any group. Gottfredson and Hirschi (1987, 1990) argue that longitudinal data analyses such as the ones undertaken in this study add nothing novel to the study of crime that cannot be determined by simply studying individuals at just one point in time.

According to the age-graded life course theory of Sampson and Laub (1993), the relationship between age and crime is due to exposure to varying levels of informal social control that individuals experience across the life course. The aggregate age-crime curve reaches it peak during adolescence because that is the period of time when the informal

social control levels (in the aggregate) are at their weakest; the social bond is under the greatest strain during this segment of the life course (see e.g., Hirschi 1969; Tittle 1988). However, the increasing forces of social control that come into play (in the aggregate) with the salient life events of adulthood serve to reduce criminal activity throughout adulthood. More importantly, Sampson and Laub foresee change in behavior as possible for all offenders (at the individual-level), whether they are of high or low criminal propensity. The opportunity for change is available for offenders, even though some offenders may not experience any change at all—the change may come at later ages for some offenders compared to others, or it may be negative for some offenders (i.e., increasing arrest rates) and positive for others (i.e., desistence). The notion that change is possible for all individuals (i.e., for any offender type), regardless of the prior nature of their offending history, is critical for Sampson and Laub because it speaks to the position that differences in criminal propensity/criminal activity are not necessarily stable across time. In fact, Sampson and Laub's theory, which focuses largely on the salient life events of adulthood (e.g., marriage, family, employment), stipulates adulthood as the precise period when the preexisting differences between individuals become less relevant for crime participation than the salient life events of adulthood that often produce a corresponding increase of informal social control. In other words, there should be between-individual changes in the rates of criminal activity during adulthood that cannot be explained as the mere unfolding of preexisting differences carrying over from the early childhood years. Sampson and Laub, in fact, challenge Gottfredson and Hirschi's "invariance" argument because: (1) it cuts at the core conceptual foundations of the life course perspective, especially the presumption that time and place matter in the lives of

individuals; and (2) because of the ontogenetic fallacy overtones of their invariance argument (i.e., attributing crime as due solely as a consequence of a preexisting personal trait of the individual) that relegates the experiences of adulthood (and their supposed benefits) to mere selection effects. Clearly, longitudinal data/analyses are critical to the theory of Sampson and Lab (1993).

Finally, the dual taxonomy theoretical exposition of Moffit (1993) predicts both continuity and change in criminal offending across time within populations, although the words "continuity" and "change" each apply to separate theoretical offender groups identified by Moffit. Change (discontinuity) defines the offending patterns of the "adolescent-limited" offender group, whereas stability (continuity) defines the offending patterns of the "life-course-persistent" group. Important for the topic of this chapter is Moffitt's (1993, 1997) argument that the mixture of the two hypothesized offender types makes the aggregate age-crime curve assume its observed shape. The rapid upswing in rates of criminal activity during adolescence results from the increasing participation rates of the adolescent-limited offender group, whereas the downward swing results from the patterns of desistance in adulthood of this group.[2] Given that the adolescent-limited group is assumed to be the more prevalent (of the two offender groups), the offending patterns of the adolescent-limited group are argued to determine the overall shape of the curve, while the life-course-persistent offenders are largely responsible for composing the childhood and adulthood tails of the age-crime curve. The key assumption of this

[2] This argument is similar in tone to that given by Blumstein and his colleagues (Blumstein and Cohen 1979, 1987; Blumstein et al. 1986, 1988a, 1988b; Farrington 1983, 1986). The difference is that Blumstein et al. argued that all offenders commit crimes at a constant rate as long as they are *active* offenders and that the aggregate age-crime curve assumes its shape as a result of varying levels of participation (prevalence) rates across age.

theoretical explanation is that it is empirically possible to not only distinguish between the two different groups (with longitudinal data covering the childhood through adulthood years), but also that it is only necessary to identify two distinct offender groups in the offender population (Fergusson et al. 2000). Similar to Sampson and Laub's (1993) theory, the need for longitudinal data is also crucial for demonstrating the empirical adequacy of the dual taxonomy theory of Moffitt (1993).

Chapter 3 of this study reviewed the extant literature on studies of the age-crime curve (within homogenous latent classes of offenders). The review discussed several current limitations with studies on this topic that require further research. There were two key limitations to prior studies discussed in Chapter 3. First, only one study to date (by D'Unger et al. 1998) has examined data from more than one dataset generalizable to the same population over time. This relative paucity of research has made it difficult not only to replicate the existence of a crime trajectory group over time, but also to establish whether there are any changes in the precise number or nature of the offending trajectories over time. The second main limitation noted in Chapter 3 is that, to date, there have been only two studies of the age-crime curve that examine samples of "high-risk" offenders (Laub et al. 1998; Piquero et al. 2001), and that each of these studies has limitations (e.g., use of nonrandom samples, samples limited to white juveniles only, data gathered in the 1930s, limited segment of the age distribution studied) that leave several vital questions concerning the relationship between age and crime with this population unanswered. Key questions we previously noted were: (1) how many "latent classes" of offenders are necessary and sufficient to capture the variation of offending trajectories in the serious youthful offender population; (2) how do differences in offending trajectories

during the juvenile years relate to the nature of offending during the adult years (i.e., are the between-group differences maintained over time); (3) is there an adolescent-peaked group within this population?

Accordingly, we now present a series of analyses in this chapter that address these questions using the three samples of (very) "high-risk" offenders from the CYA. This chapter has six main sections. In the first section, we present a descriptive summary of the aggregate age-crime relationship within the three samples. In the second section, we describe the process through which we arrived at the optimal number of latent classes for each sample. The optimal number of latent classes will provide evidence regarding the empirical support in these data for hypotheses $\mathbf{H}_1$ and $\mathbf{H}_2$ described in Chapter 3. The third section contains a statistical test of Gottfredson and Hirschi's hypothesis that there is uniformity in the shape of the age-crime curves of the latent classes by testing the statistical significance of allowing the magnitude of the age parameters to vary across the latent classes (hypothesis $\mathbf{H}_4$ in Chapter 3). The final three sections present substantive evidence concerning the hypotheses for the three samples.

As discussed in Chapter 3, the results presented in this chapter are derived from use of the semiparametric mixed Poisson model developed by Nagin and Land (1993). Using the final analytic samples constructed for each of the three CYA samples, the following model was estimated on each of the three samples:

$$\ln\left(\lambda_{it}^{j}\right) = \left(\beta_0 + \tilde{\varepsilon}_j\right) + \left((age_{it}/10) * \beta_{age}^{j}\right) + \left(\left(\left(age_{it}^2\right)/100\right) * \beta_{age^2}^{j}\right). \qquad (1)$$

Models allowing up to 8 latent classes were tested, and the BIC statistic along with empirical testing assessing whether the solutions are global maximums was used to select the optimal number of latent classes or components in the mixture distribution (discussed in further detail below). The finite mixture models employed in this study were estimated using the software program Latent GOLD, Version 2.0.9 (Vermunt and Magidson 2001).

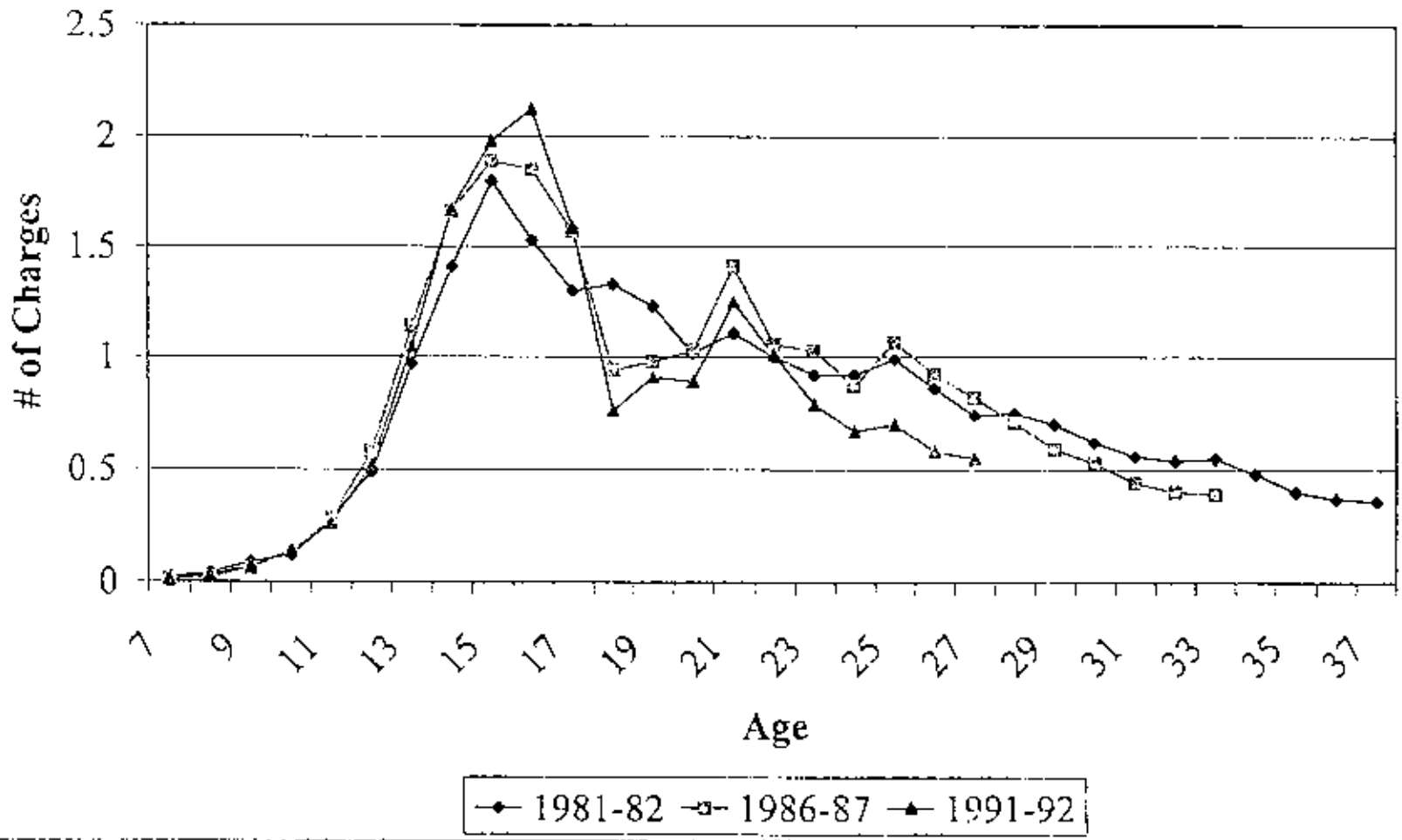
DESCRIPTIVE SUMMARY OF THE AGE-CRIME RELATIONSHIP

Prior to examining the heterogeneity with respect to the age-crime trajectories that exist within the serious youthful offender population, we first present an overall "baseline" description of the age-crime relationship. Figure 7.1 presents the aggregate age-crime curve in the three samples.[3]

As seen in Figure 7.1, the overall shape replicates the robust shape of the aggregate age-crime curve, with rapidly increasing arrest rates in adolescence, a peak arrest rate during adolescence, and then a decline in the arrest rate through adulthood. It is informative to note that the trend of the age-crime curves is amazingly similar across the three samples, which is the type of finding that originally sparked the "invariance" argument of Gottfredson and Hirschi. In fact, the three samples are nearly identical up through about age 15, after which there is more variability from sample to sample; still

[3] We have only graphed the curves up through the last age at which at least 50% of the sample was available (at-risk) for estimation. Again, a description of the percentage of cases that were at-risk at each age is presented in Appendix C.

Figure 7.1. Aggregate Age-Crime Curve, by Sample



the overall pattern is relatively robust from sample to sample.[4] The peak age of arrest in these three samples actually peaks earlier than the typical aggregate age-crime curve computed using overall population-level data (as presented in Figure 1.1), but this is not surprising given that the samples are composed of known youthful offenders who by definition were active at an early age in life. For example, in the 1981-82 sample, the peak age of arrest was age 15 (1.8 arrest charges). The peak age was also age 15 in the 1986-87 sample (1.89 arrest charges), and the peak age was 16 in the 1991-92 sample (2.12 arrest charges). In the final age presented in Figure 7.1 (which is the last age at which at least 50% of the sample was available), the average number of arrest charges was 0.36 at age 37 (1981-82 sample), 0.33 at age 33 (1986-87 sample), and 0.55 at age 27 (1991-92 sample). It is interesting to note that the 1991-92 sample experienced a more rapid decline in their arrest rates compared to the other two samples, and this is the sample that would have been most likely to have benefited from the crime rate and corresponding arrest rate decline of the mid to late 1990s. Examining this issue (i.e., how possible "period effects" affect serious offenders is an interesting question), however, it is far beyond the scope of this study. Thus, in the aggregate, the age-crime curve for these three samples resembles the overall aggregate age-crime curve (albeit with an earlier

[4] The dip in the rate between the ages of 18 and 20 for the 1986-87 sample and the 1991-92 sample probably: (1) partially reflect the fact that only ages during which the offenders were incarcerated in the CYA for the *entire* "age year" during sample stay were excluded from being at-risk; and (2) probably partially reflecting the increased parole revocation rates for those two samples compared to the 1981-82 sample. Unfortunately, determining, the precise cause of the dip is impossible because we do not have access on the incarcerations in the CYA for parole revocations. If a ward's parole was revoked after the date of release on the sample stay, we did not have access to the exact dates of the subsequent incarcerations in the CYA for the parole revocation, and thus some individuals would appear to be at-risk when they really were incarcerated in the CYA for the entire age-year. It is probably the joint interaction between those two possibilities that is largely responsible for the drop. It is our contention that the drop is artifactually the result of complex at-risk processes, especially when you note the trend is back "on track" at age 20

peak). Our attention now turns to examining whether there are diverse, heterogeneous age-crime trajectories among different "offender types" (or latent classes) concealed or masked within the overall age-crime curve (computed for all of the sample cases combined). In the next section, we discuss the selection of the optimal number of latent classes.

SELECTING THE OPTIMAL NUMBER OF LATENT CLASSES

As described in Chapter 5, models with varying numbers of points of support are not nested within each other, and thus alternative methods (besides the likelihood ratio test) of determining the number of mixture components have been recommended. Following the lead of D'Unger et al. (1998) and Nagin (1999), we used the Bayesian Information Criterion (BIC) statistic for selecting the appropriate number of groups in the mixing distribution.

However, it was also pointed out in Chapter 5 that the selection of the optimal number of points of supports for the mixing distribution is complicated by the fact that mixture models are known to often have a problem with "local solutions" (i.e., a unique global solution is not always achieved) (Goodman 1974; Cameron and Trivedi 1998; Wedel and Kamakura 1998; Vermunt and Magidson 2001). This issue concerns whether the likelihood function is unimodal (i.e., globally concave) or multimodal. In brief, the local solution problem arises from the fact that it is possible to for the model's numerical algorithm to converge to multiple local maxima. Following the recommendations of Wedel and Kamakura (1998) Cameron and Trivedi (1998), Greene (2000), and Vermunt and Magidson (2001), we tested extensively for the presence of local solutions through

the use of "multiple sets of starting values" coupled with multiple (10) runs of the estimation procedure. We describe this local solution problem and the testing process below, but first we briefly describe the method of estimation employed in generating the results presented herein.

METHOD OF ESTIMATION & GLOBAL/LOCAL SOLUTIONS

The parameters for the semiparametric mixed Poisson models were obtained via a maximum likelihood (ML) estimation routine employing the Newton-Raphson algorithm.[5] The ML estimation routine that maximizes the likelihood function is

$$\ln L = \sum_i n_i \ln f\left(\mathbf{y}_i | \mathbf{X}_i, \boldsymbol{\theta}\right), \qquad (2)$$

where $\boldsymbol{\theta}$ is a vector of estimated model parameters, $\mathbf{y}_i$ and $\mathbf{X}_i$ are vectors of dependent and independent variables, respectively, for the ith vector pattern, and n_i is the number of cases with the given i^{th} vector pattern.[6]

In equation (2), $f\left(\mathbf{y}_i | \mathbf{X}_i, \boldsymbol{\theta}\right)$ is the Poisson probability density function for covariate pattern i given the estimated parameter values contained in the vector $\boldsymbol{\theta}$. The

[5] Further detailed information on using maximum likelihood (ML) estimation methods for finite mixture (latent class) models is available in Haberman (1988) and Vermunt and Magidson (2001), and the maximum likelihood method is described generally in Greene (2000). This presentation draws heavily on the exposition of ML methods in Vermunt and Magidson (2001).

[6] The computation of the estimates of the semiparametric mixed Poisson model is simplified by using "frequency weighted" data where individuals are grouped together according to a shared vector pattern because it reduces the dimensionality of the matrices employed during the estimation process. That is the purpose for the n_i covariate vector pattern component in equation (2). For example, if 1,000 individuals remain arrest free at age 7, those individuals can represented by 1 row of data that is "frequency weighted" by 1000, rather than incorporating the information through 1,000 rows.

estimation task is to find a set the estimates, denoted as $\hat{\theta}$, for the parameter vector θ such that partial derivatives of the log-likelihood function with respect to θ are equal to zero:

$$\frac{\partial \ln L}{\partial \hat{\theta}} = 0. \tag{3}$$

In this study, the estimates were obtained using the Newton-Raphson (NR) algorithm, which is a general numerical optimization method. The NR method is an iterative, gradient-based method of optimization (Greene 2000):

$$\theta^{\nu} = \theta^{\nu-1} - \varepsilon \mathbf{H}^{-1}\mathbf{g}, \tag{4}$$

where ν denotes the current j^{th} iteration, $\mathbf{g}$ is the gradient vector containing the first-order derivatives of the log-likelihood function evaluated at $\theta^{\nu-1}$ (i.e., at the prior iteration of the algorithm), ε is a scalar number known as the "step" size, and $\mathbf{H}$ is the Hessian matrix (or observed information matrix) that contains the second-order derivatives. The elements of the gradient vector, denoted as g_r, are equal to

$$g_r = \sum_i n_i \frac{\partial \ln f(\mathbf{y}_i|\mathbf{X}_i, \theta)}{\partial \theta_r}, \tag{5}$$

and the elements of the Hessian matrix are:

$$H_{rs} = \sum_i n_i \frac{\partial^2 \ln f(\mathbf{y}_i | \mathbf{X}_i, \boldsymbol{\theta})}{\partial \theta_r \partial \theta_s'} . \tag{6}$$

The asymptotic variance-covariance matrix of the maximum likelihood estimates, which are the final estimates $\hat{\boldsymbol{\theta}}$ that solve equation (3) above, are calculated by inverting the Hessian matrix (the negative of the matrix of second derivatives of the log-likelihood function with respect to the estimated parameters)

$$V(\boldsymbol{\theta}) = \mathbf{H}^{-1} = -\left[\frac{\partial^2 \ln L(\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}} \, \partial \hat{\boldsymbol{\theta}}'} \right]^{-1} . \tag{7}$$

Finally, *convergence* of the algorithm occurs when the sum of the absolute relative changes in the parameter estimates, or

$$\sum_r \left| \frac{\hat{\theta}_r^{\nu} - \hat{\theta}_r^{\nu-1}}{\hat{\theta}_r^{\nu-1}} \right| , \tag{8}$$

meets a prior defined tolerance criterion whereby the absolute change in the parameters meets the convergence criterion. In this study, we employed the use of a very "tight" convergence criterion: 1e-10 (or 0.0000000001).[7]

[7] Ideally one would like to terminate the estimation algorithm when the gradient is actually zero; however, in practice, such a convergence criterion is problematic because of the accumulated rounding error that

The local solution problem arises because the condition defined in equation (3), can often be satisfied with several different solutions such that a minor change in any one of the estimated parameters leads to a degradation of the log-likelihood value. The algorithm only distinguishes that a maximum has been reached, but it cannot determine whether the solution is a global or local maximum solution.[8] The distinction between the two is that a local solution is the best solution in a particular *area* of the parameter space, whereas a global solution is the best solution in *the* parameter space.

Perhaps a non-statistical description will help clarify this problem. Imagine that you have to climb a mountain blindfolded ten times and you have to place a flag at the top of the mountain when you think you are at the top (i.e., maximized the likelihood function).[9] You climb the mountain using some pre-defined "step" sizes that takes you, according to your exact step rule (i.e., the numerical algorithm), up the mountain to a specific location. If you started climbing from the same spot every time you climbed the mountain (i.e., use the same starting values), you will always end up at the exact same spot because you take "steps" according to a very specific rule. Your ten flags will all be at the same spot simply because that is where your "step rules" tell you the top of the mountain is located, not because it actually is the top of the mountain. If however, you started your climb up the mountain from a different location each time (i.e., use different

---

results largely as a consequence of using digital computers employing the use of binary mathematics (Greene 2000).

[8] To date, the issue of local solutions has largely been ignored in the field of criminology, primarily because the only way to even test for the presence of local solutions is by estimating the model multiple times with random sets of starting values. None of the articles in criminology that employ the use of the finite mixture model discuss the issue of local solutions, much less actually test for the presence of them. Perhaps our results are anomalous, but they do indicate that local solutions may be a serious problem with the high-order latent class models (e.g., 7-class, 8-class models).

[9] Of course, this is an oversimplification of the problem because of the actual dimensionality of the likelihood function, but the imagery accurately conveys the nature of the problem.

starting estimates) and you ended up at the same location each time regardless of where you started your ascent (i.e., same likelihood value and same model parameter estimates regardless of the value of the starting estimates), you would be more confident that you had actually reached the top of the mountain (i.e., attained a global solution). If on the other hand, you placed your flags at multiple sites (i.e., attained multiple local solutions), you would not be able to tell if you had actually reached the top of the mountain, nor would you be confident that you really know where the top of the mountain is located (i.e., what the true global solution is assuming there is one). The only thing you would know was that at certain sites on the mountain peak, your "step" size calculations indicated to you that if you took any other steps, you would start going "down the mountain" again (i.e., a degradation of the log-likelihood value). Of course, the actual peak of the mountain may be across the valley you were starting down into when your step size rule told you to stop. But since you cannot see the valley and the peak of the mountain across the valley (you're blindfolded), you simply think you are at the top of the mountain and stop. Our approach to the possible problem of local solutions is to climb the "likelihood mountain" several times and see if we land at the same location each and every time.

The testing for global/local solutions process was as follows. Initially, 50 sets of random starting estimates were generated. These 50 sets of starting estimates were then run through 20 iterations of the EM algorithm.[10] At this point, the top 10% of the starting

[10] The EM (expectation maximization) algorithm (Dempster et al. 1977) was used in this step of the process because the EM algorithm is extremely stable and fast when the model is far away from the optimum, whereas the NR algorithm is inefficient and computationally burdensome (because of the calculation of the Hessian matrix) when far away from the optimum, but is extremely efficient and accurate when near the optimum (Wong et al. 1998; Greene 2000; Vermunt and Magidson 2001). Thus, the method employed here

seeds with the largest log-likelihoods at this point were then run through an additional 40 iterations of the EM algorithm. Finally, the set of starting estimates with the largest log-likelihood value at this point was then entered into the ML algorithm, which either ended in convergence (i.e., a solution) or the maximum number of iterations (set at 150) was reached without convergence and algorithm stoops. This entire process was repeated ten times for each point of support model (except for the 1 point of support model which was only run once).[11] Thus, the 2 point of support model was run ten times, the 3 point of support model was run 10 times, and so on. The final solutions of the models were then compared to evaluate if the same unique solution was generated with each run of the model. Models with more than one solution (i.e., local solutions) were identified through this extensive model testing procedure.

THE OPTIMAL NUMBER OF LATENT CLASSES IN THE THREE SAMPLES

To determine the optimal number of latent classes, we jointly used the BIC statistic and the testing for the local/global solutions described above. The rule, in short, was the model that resulted in the largest (least negative) BIC value in the presence of

---

of using the EM algorithm in the initial stages and then switching to the NR algorithm takes advantages of the strengths of each algorithm (Wang et al. 1998; Vermunt and Magidson 2001). A further advantage of the NR method is that it also provides estimates of the standard errors of the estimated parameters, using equation (7), based on the final Hessian matrix (from the final iteration of the NR algorithm).

[11] One might wonder why we chose to run the models 10 separate times. Our initial testing of the problem indicated the 5 separate runs did not always identify local solutions, but 10 separate runs caught all of the local solutions we encountered. The computational burden of even running the models 10 times was quite severe, especially since this entire process was repeated three separate times (once for each sample). Because the 1981-82 sample had the most data points, this sample was the most computationally burdensome. Each run of the estimation procedure for this sample (i.e., estimating a 1-class through 8-class model) lasted anywhere from 6 to 8 hours on a 1300mhz AMD Athlon with 512 MB RAM. Estimating the models 10 times on this sample required just over 75 hours of actual computer time in total, which was magnified by the fact Latent GOLD has no batch mode version of running the program and collecting the output. In total, the actual computer time for estimating the models 10 times on each of the three samples was just under 200 hours. This entire process had to repeated all over again for the models presented in Chapter 7 because they involved a different specification.

only a single solution would be chosen as the optimum number of latent classes. Models with more than one (local) solution were not considered as possible models. This rule turned out to be highly effective in choosing a model.

The results of fitting the 1-class through 8-class semiparametric mixed Poisson models for all three samples, according to the specification of equation (1), are presented in Table 7.1. The results of the local/global testing indicated that in all three samples the 7-class and 8-class models were prone to multiple local maxima solutions that varied from one solution to another.[12] In all three samples, however, the 6-class models generated the same unique solution all ten times they were estimated. Further, and as shown in Table 7.1, the 6-class model in all three samples had the largest BIC value (note the large positive values in the "Change in BIC" column of Table 7.1 for the 6-class models), and thus the 6-class models were determined to be the models with the optimal number of latent classes for each of the three samples. Note that the addition of a sixth point of support increased the BIC value by 427.84 in the 1981-82 sample, 190.66 in the 1986-87 sample, and 183.75 in the 1991-92 sample.

The results presented so far allow for an examination of the empirical support for the hypotheses $\mathbf{H_1}$ and $\mathbf{H_2}$ delineated in Chapter 3. The empirical evidence in this study (for all three samples) supports the hypothesis, $\mathbf{H_1}$, that there are multiple, distinct offender groups on the high-end of the criminal propensity continuum. These results support the contentions of Cohen and Vila (1996) and D'Unger et al. (1998) that the far end of the continuum appears to have far greater heterogeneity than previously thought.

[12] Examination of the estimates from the local solutions indicates that these models were "weakly identified," meaning that two or more of the points of support were not very different from one another. See footnote 6 of Chapter 8 for an extensive discussion of weak identification.

Table 7.1. BIC Statistics for the Semiparametric Mixed Poisson Models, by Sample

# of Latent Classes	Log-Likelihood	BIC	Change in BIC
1981-82 Sample			
1	-82026.06	-82037.45	---
2	-77059.40	-77085.99	4951.46
3	-75404.98	-75446.75	1639.23
4	-74100.77	-74157.74	1289.01
5	-73509.48	-73581.63	576.10
6	-73066.44	-73153.79	427.84
7	MS	MS	---
8	MS	MS	---
1986-87 Sample			
1	-52620.27	-52631.18	---
2	-49806.40	-49831.86	2799.32
3	-49024.08	-49064.09	767.78
4	-48347.17	-48401.73	662.35
5	-48032.93	-48102.03	299.70
6	-47827.72	-47911.38	190.66
7	MS	MS	---
8	MS	MS	---
1991-92 Sample			
1	-39923.26	-39934.16	---
2	-37952.46	-37977.89	1956.27
3	-37376.58	-37416.55	561.34
4	-36984.44	-37038.95	377.60
5	-36748.70	-36817.75	221.20
6	-36550.41	-36634.00	183.75
7	MS	MS	---
8	MS	MS	---

Note: "MS" indicates multiple solutions were found, with no clear global solution.
The model that has the least negative value of the BIC statistic is the favored model.

Empirically, the results presented so far also shed considerable doubt on Moffitt's hypothesis (H_2) that there are only two discrete offender groups within even just the serious youthful offender population.[13] The BIC statistics in all three samples favored a 6-class model, and the positive increase in the magnitude of the BIC statistics between the models that only allow for two latent classes (corresponding to the hypothesized number of discrete groups in the Moffitt theory) and those that allow for six latent classes were indeed quite large (3932.2 in the 1981-82 sample; 1920.5 in the 1986-87 sample; and 1343.9 in the 1991-92 sample). It is important to note that Keribin (1997) and Wang et al. (1998) both demonstrate that that the BIC statistic identifies the optimal number of components (points of support) in the mixing distribution with a high degree of precision.

The resulting parameter estimates for the 6-class models are presented in Table 7.2. Due to the fact this model is full of nonlinear terms, it is difficult to substantively interpret the parameter estimates of this model (as presented in Table 7.2). Below, we present a substantive presentation of each model's estimates by graphically depicting the offending trajectories, but for now we make several comments regarding these parameter estimates. First, it is apparent from the signs of the age and age-squared coefficients of each latent class that there was a significant nonlinear relationship between age and crime for all 6 latent classes in each of the samples. There was no latent class in any of the samples that was found to offend at a relatively constant rate across age in the "spirit" of the life-course-persistent offender group. The age and age-squared parameter estimates for all of the latent classes were significantly different from zero (and indicative of a

[13] Importantly, regardless of whether one assumes a 3-class, 4-class, 5-class, or 6-class model, the evidence rejects the notion that there are only two latent classes or offender typologies present in the population; this appears to be the case even here among the three random (but select) samples of serious youthful offenders

Table 7.2. Parameter Estimates From the 6-Class Semiparamteric Mixed Poisson Model, by Sample

	Sample		
	1981-82	1986-87	1991-92
Random Effects			
Class 1	-4.860	-8.473	-7.346
	(36.06)	(28.28)	(24.10)
Class 2	-3.057	-5.062	-17.921
	(34.18)	(27.02)	(16.67)
Class 3	-5.771	-3.378	-8.321
	(33.60)	(16.52)	(32.57)
Class 4	-8.624	-24.448	-5.877
	(16.54)	(18.57)	(16.48)
Class 5	-22.006	-7.073	-39.014
	(23.36)	(25.84)	(17.57)
Class 6	-4.000	-4.082	-4.887
	(29.71)	(22.35)	(17.11)
Age Effects			
Class 1 Age	3.989	9.6355	7.554
	(33.53)	(28.36)	(21.16)
Class 1 Age Sq.	-0.813	-2.6269	-1.905
	(31.47)	(27.36)	(18.85)
Class 2 Age	3.651	4.5456	21.464
	(41.34)	(26.71)	(16.56)
Class 2 Age Sq.	-0.884	-0.9924	-6.217
	(41.87)	(25.55)	(15.95)
Class 3 Age	6.870	4.5163	9.876
	(35.45)	(19.06)	(34.78)
Class 3 Age Sq.	-1.881	-1.2677	-2.603
	(35.19)	(18.98)	(33.55)
Class 4 Age	7.691	32.0432	8.141
	(13.90)	(18.24)	(19.12)
Class 4 Age Sq.	-1.882	-10.2455	-2.553
	(13.02)	(17.64)	(20.55)
Class 5 Age	29.199	7.9416	52.885
	(23.39)	(28.31)	(17.40)
Class 5 Age Sq.	-9.375	-1.9275	-17.518
	(22.86)	(28.11)	(16.92)
Class 6 Age	3.592	3.7658	4.753
	(31.73)	(22.89)	(16.25)
Class 6 Age Sq.	-0.666	-0.7319	-1.004
	(28.96)	(20.09)	(13.81)
N (Observations)	1989	1443	1434
N (Panel Observations)	60453	37390	29385

Note: Absolute values of t-statistics are in parentheses.

significant quadratic relationship). Moffitt (1993: 695) explicitly argues that age is not a predictor of the offending trajectory of the "life-course-persistent" group. These results are highly inconsistent with the notion that there is a group of offenders whose offending behavior is not significantly variable over the age distribution.

STATISTICALLY TESTING THE AGE INVARIANCE HYPOTHESIS

In this section, we present results from statistical tests of the age invariance hypothesis. As depicted in Table 7.2, the nature of the age coefficients seem to be highly variable over the latent classes—this drives at the core of the Gottfredson and Hirschi "age invariance" hypothesis. It does not appear from the results presented in Table 7.2 (in any of the three samples) that the age and age-squared variables in each latent class are hovering in the neighborhood of a common value shared by all of the latent classes. The nature of the relationship between age and crime appears to be highly variable across the latent classes. For example, the estimate of the age parameter in the first latent class in the 1981-82 sample was 3.99, whereas the estimate of the age parameter in the fifth latent class was 29.20. Similar discrepancies in the magnitude of the age coefficients can be found in both the 1986-87 and the 1991-92 samples. The results presented thus far are inconsistent with the hypothesis of Gottfredson and Hirschi that there is uniformity in the shape of the age-crime curves across the six latent classes (hypothesis H_4 in Chapter 3), however, a more formal statistical test is required to adequately judge the empirical validity of the hypothesis of Gottfredson and Hirschi. We will now formally test this hypothesis.

Recall that Gottfredson and Hirschi (1990) argue in favor of a robust relationship between age and crime—while groups differ on their mean offending rate, the actual shapes of the trajectories are identical. In other words, the shapes of the age-crime curves are invariant. This predicted relationship corresponds to a specification of the finite mixture model that allows the latent classes to have intercept values that vary over the latent classes, but the age parameters (i.e., age and age squared) must share a common (invariant) value. This hypothesis can be evaluated without re-estimating the model by using the Wald statistic to test the equality of each set of age coefficients across the latent classes via a linear constraint (Long 1997; Powers and Xie 2000; Vermunt and Magidson 2001). The Wald statistical test that the age coefficients are identical across the latent classes tests the null hypothesis

$$H_0 : \mathbf{Q}\boldsymbol{\beta} = \mathbf{r}, \tag{9}$$

where $\boldsymbol{\beta}$ is a vector of regression parameters being tested, $\mathbf{Q}$ is a matrix of constraints, and $\mathbf{r}$ is a vector of constants (Long 1997). The null hypothesis of equation (9) can be tested using the Wald statistic that is calculated as:

$$W = \left(\mathbf{Q}'\hat{\boldsymbol{\beta}} - \mathbf{r}\right)'\left(\mathbf{Q}\hat{\mathbf{H}}^{-1}\mathbf{Q}'\right)^{-1}\left(\mathbf{Q}'\hat{\boldsymbol{\beta}} - \mathbf{r}\right), \tag{10}$$

where H is the Hessian matrix with elements described above in equation (6). The Wald statistic is chi-square distributed with degrees of freedom equal to the number of imposed constraints (Long 1997).

Table 7.3 presents the results generated from calculating the Wald statistic testing the constraint that the parameter estimates for (1) the intercept, (2) the age variable, and (3) the age-squared variable were equal across the six latent classes. As shown in Table 7.3, the hypotheses that the age and age-squared parameter estimates were equal across all six latent classes were resoundingly rejected in all three samples. We say "resoundingly" because of the size and significance of the Wald statistics shown in Table 7.3. For comparative purposes, note that a chi-square value of 16.75 (with 5 degree of freedom) is significant at the 0.005 level. The smallest value of Wald statistic shown in Table 7.3 for either the age or age-squared variables is 477.84 (for the age variable in the 1991-92 sample), which is extremely large compared to 16.75. Thus, the intercepts for each of the classes were found to not be equal to an overall constant value (which is consistent with the argument of Gottfredson and Hirschi). More importantly, the estimated age parameters for both the age and the age-squared variables in all three samples were found to be highly inconsistent with the hypothesis that they were all equal to the same value. Thus, in all three samples, there is strong statistical evidence that refutes the hypothesis (H_4) that the age-crime curves among the latent classes are invariant. They differ not only with respect to their mean rate, but they also differ

Table 7.3. Wald Statistics Testing the Equality of the Estimated Parameters Across the Six Latent Classes , by Sample

Paramter Tested	Wald Statistic	Degrees of Freedom	p-value
1981-82 Sample			
Intercept	533.63	5	0.0000
Age	635.07	5	0.0000
Age-squared	791.36	5	0.0000
1986-87 Sample			
Intercept	566.49	5	0.0000
Age	593.53	5	0.0000
Age-squared	737.72	5	0.0000
1991-92 Sample			
Intercept	456.94	5	0.0000
Age	477.84	5	0.0000
Age-squared	545.66	5	0.0000

(significantly) with respect to the developmental shape or nature of their age-crime curve.[14]

To this point, we have presented statistical evidence testing hypotheses $\mathbf{H_1}$, $\mathbf{H_2}$, and $\mathbf{H_4}$. The evidence testing these hypotheses supports hypothesis $\mathbf{H_1}$, but rejects hypotheses $\mathbf{H_2}$ and $\mathbf{H_4}$. However, as students of statistics learn very early, there is a difference between statistical significance and substantive importance (especially when dealing with large samples). Accordingly, we now turn to a substantive-based approach of examining these hypotheses. This substantive, graphically-based method also allows us to address hypothesis $\mathbf{H_3}$ concerning the existence of an adolescent-limited (or adolescent-peaked) offender type within these samples. The numerical focus to this point has not allowed for the determination of whether there was an adolescent-limited offender group in each or any of the three samples.

The substantive approach we use here makes use of the assignment of each individual in the sample to a particular latent class using the posterior assignment probabilities (described in Chapter 5) that indicate the probability that each latent class has generated the individual's observed longitudinal pattern of offending. This assignment of each individual to a latent class allows for a descriptive summary and

[14] For comparative purposes, we did re-estimate the semi-parametric mixed Poisson model according to the specification that constrained the estimated age parameters to be equal across the six latent classes. Comparisons of the log-likelihood values for these models compared to the log-likelihood values from the models where the estimated age parameters were allowed to vary across the latent classes indicated the same conclusion reached above—there is a significant improvement in the fit of the model when the estimated age parameters are allowed to vary across the latent classes. Calculation of the likelihood ratio test statistic indicated that there was highly significant improvement in model fit by allowing the estimate of both the age and age-squared variables to vary across the latent classes. The likelihood ratio test statistic in this situation tests the *simultaneous joint significance* of freeing both the age and age-squared variables, whereas the results presented in Table 7.3 tested the significance of constraining each of the age regression coefficients by themselves. Regardless of which test is used, the conclusion is identical across both statistical tests—both tests overwhelmingly rejected the hypothesis that the shapes of the trajectories were equal across the latent classes.

graphical depiction of how substantively different the latent classes are with respect to their observed overall rate of offending and their longitudinal offending patterns. In the next three sections (one section for each sample) we make use of these posterior assignment probabilities--each individual in the sample was "assigned" to the latent class to which the individual had the highest probability of belonging. After each individual was assigned to a particular latent class, the individual's membership in the latent class was considered unique and distinct. Then the summary characteristics and longitudinal offending trajectories depicting the relationship between age and crime within each of the latent classes were computed using both the offending histories of the individuals assigned to that latent class (i.e., observed offending characteristics) and the parameter estimates from the model (predicted offending characteristics). The results of these procedures allow for a substantive look at the differences between the latent classes in each sample, including their overall mean rate of offending (by the end of the follow-up period) and the nature of their trajectories across the age distribution.

The basic plan for the description of the results within each of the samples is as follows. First, there is a description of the percentage of each sample assigned to each latent class and a description of the posterior assignments probabilities. Next, the basic descriptive features of the offending patterns of the latent classes are discussed (e.g., age at first arrest, mean number of arrest charges). Then the observed and predicted offending trajectories across the age range are graphed and compared. The examination of the shape of the arrest trajectories will allow for a graphical examination of whether there is an adolescent-limited (or adolescent-peaked) offender group in these samples (H_3 in Chapter 3), as well as whether between-group differences are maintained across the

age range (H_4). Finally, a descriptive analysis is undertaken to examine what role, if any, adult incarceration time may have had on the decline in criminal offending among these three groups of serious youthful offenders.[15]

RESULTS FOR THE 1981-82 SAMPLE

The first section of "substantive" results presented here are for the 6-class semiparametric mixed Poisson model of the 1981-82 sample. The analyses that will follow for the 1986-87 and 1991-92 samples are procedurally mere replications of the same processes employed here for the 1981-82 sample.[16]

Latent Class Assignment Percentages and Posterior Probabilities

To begin with, we present a description of the percentages of each sample that were assigned to a given latent class and a corresponding description of the nature of the assignment probabilities (for assignment to that particular latent class) for those individuals assigned to each latent class. In the results that follow, all of the latent classes were named as LC1, LC2,..., LC6 according to their rank-ordered estimated prevalence

[15] As noted before, we only recently received the adult incarceration data concerning the individuals in these three samples; hence the incorporation of these data into the analyses herein is only descriptive. Future work will examine this question more definitively by creating more complex analytical files that essentially remove the individual from being at risk of arrest during periods of incarceration (through the use of a "street time" exposure or offset variable).

[16] Because it is not relevant to the substantive goals of this chapter, we have not presented a description of the background characteristic variables (described in Chapter 6) within each of the latent class to compare if any of the variables distinguish among the latent classes. Appendix E, however, does contain tables containing summary descriptions of the background characteristic variables within each of the latent classes. Given that these background characteristics are descriptive of the wards as of the time they were incarcerated in the CYA on the sample stay (and the wards were of varying ages at that point), and especially since the variables were measured after all of the trajectories were already in motion, these tables are presented in Appendix E merely for descriptive purposes. It is interesting to note, however, that there are few variables that have any numerically distinguishing values across the latent classes.

Table 7.4. Summary Descriptions of Latent Class Assignment Percentages and Posterior Assignment Probabilities: 1981-82 Sample

Panel A: Group Assignment Percentages

		Assigned	
Latent Class	Estimated %	%	N
LC1	26.25	26.09	519
LC2	20.49	20.26	403
LC3	20.16	20.81	414
LC4	12.22	12.27	244
LC5	10.80	10.81	215
LC6	10.09	9.75	194

Panel B: Posterior Assignment Probabilities

		Percentiles		
Latent Class	Mean	25th	50th	75th
LC1	0.92	0.88	0.99	1.00
LC2	0.91	0.86	0.98	1.00
LC3	0.88	0.82	0.97	1.00
LC4	0.89	0.84	0.97	1.00
LC5	0.93	0.94	1.00	1.00
LC6	0.92	0.92	1.00	1.00

in the sample (e.g., LC#, where the number simply represents the rank-ordered *estimated* prevalence of that latent class based on the estimates of the model). Thus, LC1 corresponds to the latent class that was estimated to have the largest percentage of the sample assigned to it.

The percentages of (estimated and actual) cases assigned to each latent class (Panel A) and a summary description of the posterior assignment probabilities (Panel B) are presented in Table 7.4. As found in Panel A, the maximum probability assignment rule resulted in the following percentage of cases being assigned to each latent class: 26% (LC1; n = 519); 20% (LC2; n = 403); 21%; (LC3; n = 414); 12%; (LC4; n = 244); 11% (LC5; n=215); and 10% (LC6; n = 194). Thus, there are no latent classes that contain an overwhelming percentage of the cases, but rather the latent class with the largest percentage of cases assigned to it (LC1) only had 26% of the sample in it. Also found in Table 7.4 (under the "Estimated %" column) is a model-based estimate of the percentage of cases that were expected to belong to each latent class; the estimate is from the theta vector described in Chapter 5. There is a considerable degree of congruence between the estimated percentage of the sample and the actual percentage of the sample that was assigned to the latent class. In fact, the largest discrepancy was less than 1% (0.65% in LC3).

Panel B contains a summary description of the posterior assignment probabilities. Each row of the panel pertains to a description of the assignment probabilities for the individuals who were assigned to that latent class. As is evident in the table, the average assignment probabilities were quite high, and the lowest average probability was only

0.88. Indeed, four of the latent classes had average assignment probabilities of 0.91 or greater. The lowest median probability value for any of the latent classes was 0.97, and thus over 50% of the sample members were assigned to the latent class that had a 0.97 probability or greater of having generated the individual's longitudinal offending pattern.

Summary Arrest Charge Information

Now we turn our attention to a descriptive summary of the arrest charge histories of the individuals assigned to each latent class. Table 7.5 summarizes the arrest histories for each latent class, including the average observed number of total arrest charges (at the end of the follow-up period on June 30, 2000), the predicted average number of arrest charges based on the estimates from the semiparametric mixed Poisson model, the observed number of serious arrest charges, and the average age at first arrest. To briefly recap the descriptive results presented earlier in Chapter 6, the "average individual" in the 1981-82 sample accumulated 22.8 arrest charges by the end of the follow-up period, accumulated 12.51 serious arrest charges, and was first arrested at an age of 14.22 years old. We have chosen to present the information in Table 7.5 rank-ordered by the mean observed number of arrest charges (from lowest to highest) because this makes the presentation of some of the other information in the table clearer.

In Table 7.5, a significant amount of between-class heterogeneity in the average number of arrest charges is obvious. The mean number of arrest charges varied from 5.59 all the way up to 44.10. LC6, which is the smallest latent class (10% of the sample), was the latent class that had with the highest average number of arrest charges. LC1 (11%), on the other hand, had the smallest average arrest charge total. Thus, there was a

Table 7.5. Summary Arrest Charge Information, by Latent Class: 1981-82 Sample

Latent Class	Mean Total Arrest Charges		Obs. Mean Serious Arrest Charges	Average Age at First Arrest
	Observed	Predicted		
LC4	5.59	5.83	3.92	17.24
LC5	11.59	11.47	6.99	13.57
LC1	18.83	18.89	11.06	15.51
LC3	20.48	20.45	12.23	13.06
LC2	36.35	35.96	19.04	12.26
LC6	44.10	43.83	20.35	14.26

difference of roughly 39 arrest charges between the latent classes with the highest and lowest arrest charge total. The largest latent class, LC1 (26%), had an average number of arrest charges (18.82) below the overall number of arrest charges for the sample (22.8). The other arrest charge totals for the remaining 3 latent classes were 11.59 (LC5; 11%), 20.48 (LC3; 21%), and 36.35 (LC2; 20%) respectively. Thus, although there are two latent classes that had averages near the average arrest charge total for the sample as a whole, clearly the average arrest charge total presented in Chapter 6 is not very representative of the arrest charge averages of most of the latent classes.

The results presented in Table 7.5 also indicate that the average number of predicted arrest charges (based on the model estimates presented in Table 7.2) for each of the latent classes was fairly accurate (in comparison with the observed average). For example, the model predicted members in the LC4 latent class would have 5.83 arrest charges, and they were observed on average to have 5.59. Similarly, the members of the LC6 latent class were predicted to have 43.8 arrest charges on average, and they were observed to have 44.1 arrest charges. Overall then, the model does a fairly accurate job at predicting the total number of arrest charges by the end of the follow-up period.

It is significant to note that for most of the six latent classes, a little over half of their arrest charges were for serious offenses. It is especially significant that in the LC6 latent class, the individuals in this sample were arrested on average for over 20 serious arrest charges. Thus, it is not simply the case that this is a latent class composed of chronic, non-serious offenders (all of these offenders are regarded by most people as very serious offenders). Nonetheless, roughly 50-60% of the observed arrest charges in each latent class were for serious offenses. The latent class with the highest percentage of

serious offenses was the LC4 latent class (who had the lowest number of charges), with roughly 70% of their arrest charges having been composed of serious offenses.

Finally, it is significant to note several items about the average age at first arrest across the latent classes. First, there is a considerable degree of heterogeneity between the classes with respect to their average age of onset (as measured by their first arrest). LC2 had the youngest average age of onset (12.26 years old), whereas LC4 had the oldest average age of onset (17.24). Thus, the latent class with the oldest average age at first arrest was also the group with the lowest average arrest charge total. Second, there is little semblance, however, to the rank ordering of average age at first arrest with respect to the average number of arrest charges (after noting that the latent class with the oldest average age had the lowest arrest charge total). The latent class with an average age at first arrest exactly equal to the overall average for the sample as a whole was the latent class with the highest average arrest charge total. One of the classes with an average age of onset that ranked third in terms of the youngest average age (and right around the onset of adolescence) was the latent class that had the second lowest number of arrest charges.

The reader will recall from the discussions in Chapters 2 and 6 that age of onset is a critical variable in criminological research because it is often used as a proxy variable for criminal propensity in that individuals with a younger age of onset *tend* to have higher rates of criminal activity throughout life. The fact that the rank orderings of the age of onset variable are not the reverse-identical of the average number of arrest charges is already a possible indication of a significant amount of between group differences in their trajectories of arrest. They should be reverse-identical if between group differences were stable across time because if one group is offending earlier, they should also be offending

later and at a higher rate if the between group differences are going to persist across time. Of course, these overall averages are merely suggestive of a possible problem with the "invariance hypothesis," but they do indicate possible inconsistencies with the stability of group differences hypothesis. Given the statistical results presented earlier that tested (and rejected) the hypothesis that the estimated age parameters were equal, this is the precise type of finding one would have expected.

In the next section of results we present the predicted and observed trajectories for each of the latent classes to more conclusively address and examine the issue of the stability of between-group differences across the age range. From the graphical depiction of the nature of the arrest trajectories of each of the latent classes, the question of the magnitude of the substantive differences in the nature of the growth and offending trajectories should become much more apparent.

Predicted and Observed Trajectories of Arrest

In this section, We present two figures: Figure 7.2 contains the predicted arrest trajectories (on the basis of the model parameters) for the six latent classes and Figure 7.3 contains the actual observed arrest trajectories of the 1981-82 sample.[17] The observed arrest trajectories presented in Figure 7.3 were generated by averaging the arrest charge totals at each age for all of the individuals assigned to the particular latent class that were at risk at that age.

[17] Similar to Figure 7.1, we have only graphed the curves up through the last age at which at least 50% of the sample was available (at-risk) for estimation. The age-crime curves for the over all observed ages are available in Appendix F for all three samples.

Figure 7.2. Predicted Arrest Trajectories: 6-Class Model of the 1981-82 Sample

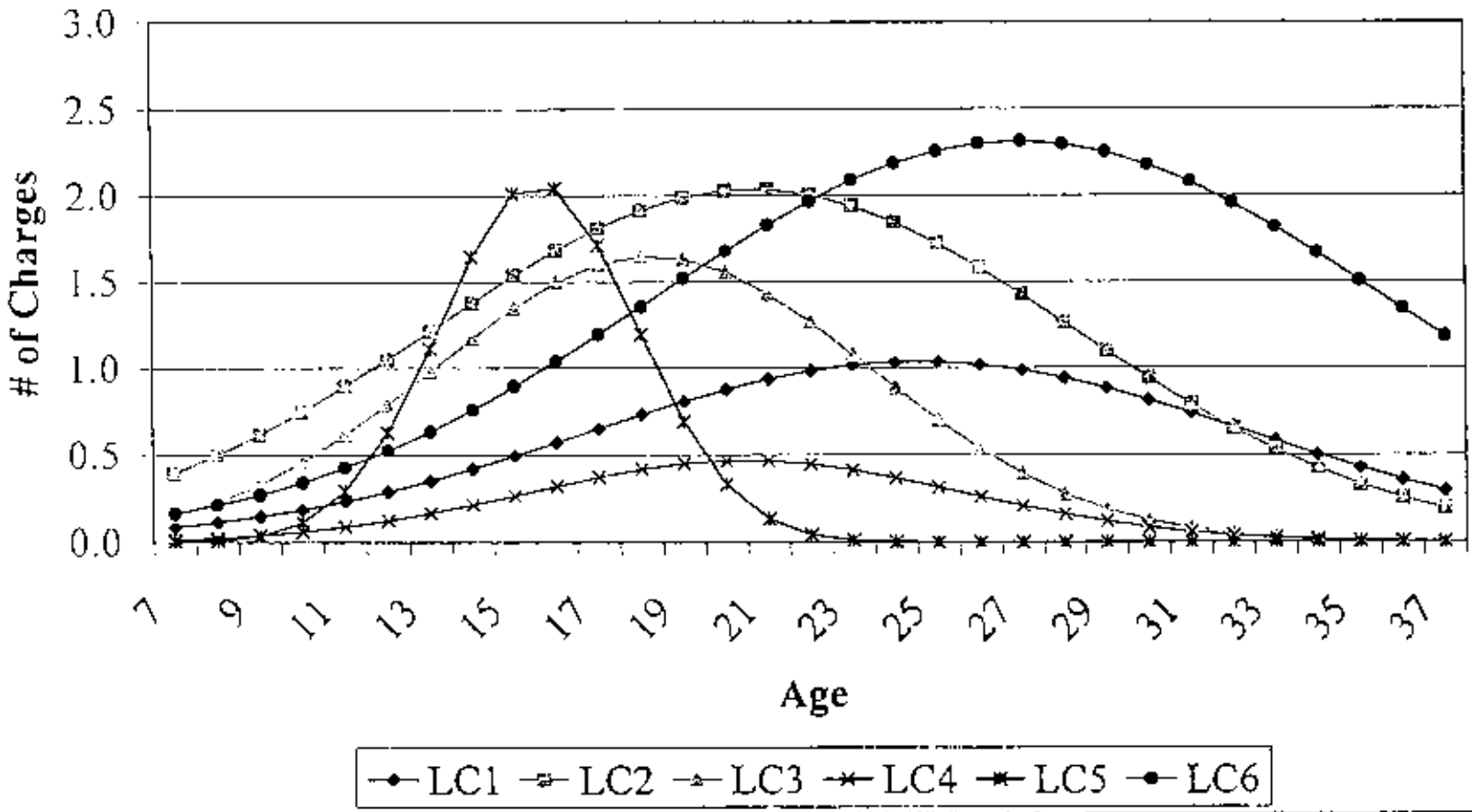
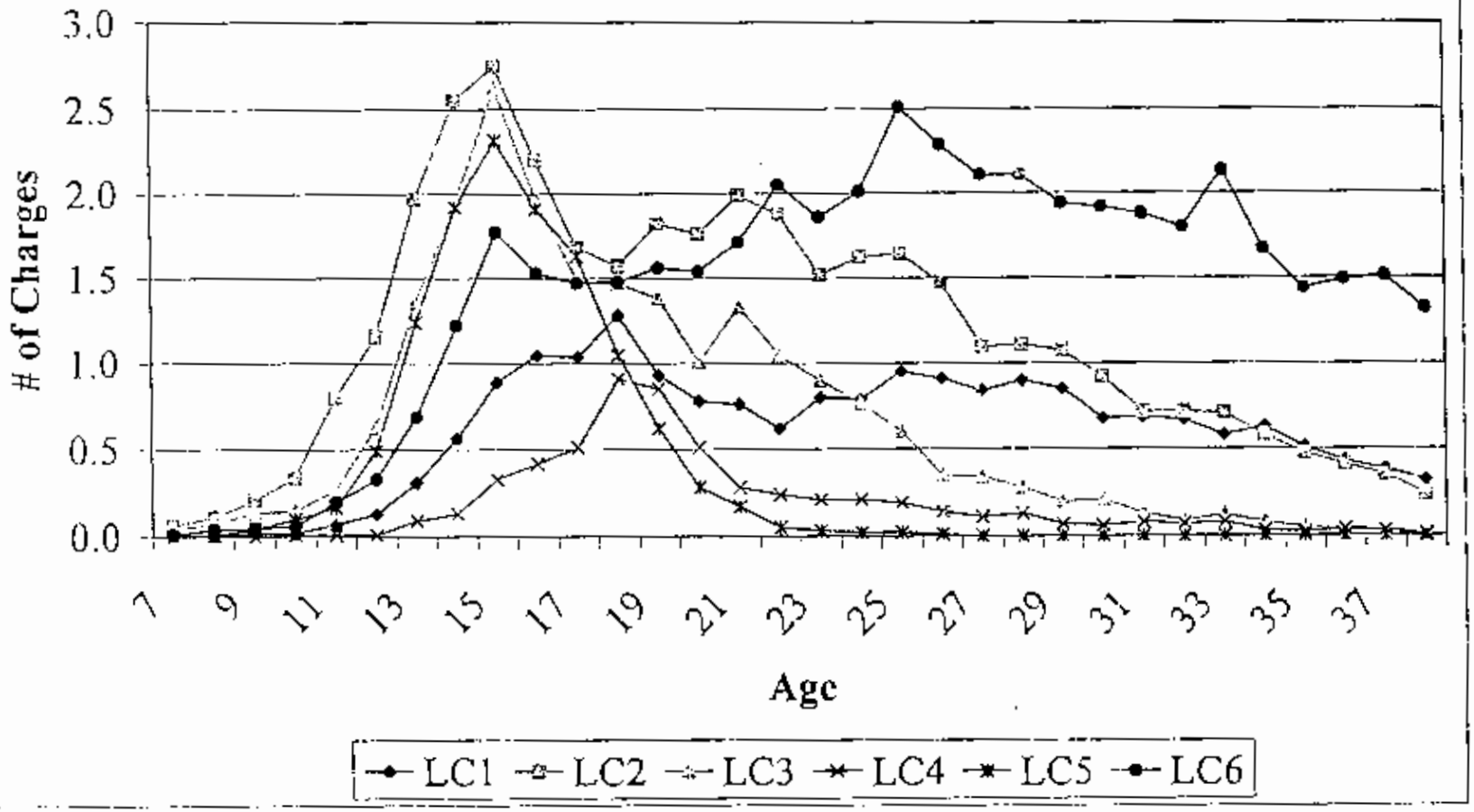


Figure 7.3. Observed Arrest Trajectories: 6-Class Model of the 1981-82 Sample



On the whole, we note that for the 1981-82 sample, the predicted trajectories tend to slightly over-estimate the arrest rates at the younger ages, undercut the peak ages, and slightly over-predict during the early 20s when the dip in the arrest rates occur (as described above in footnote 3).[18] Again, the reason for the dip *in the early twenties* is most probably an artifact of a drop in "at-risk" time due to incarceration (e.g., for both new offenses and parole revocations) during this short period of time. The average observed number of charges are back "on track" with the predicted number of charges by the mid-twenties. In sum, however, the predicted trajectories do a fairly accurate job of tracking the arrest charges at each age for each of the latent classes, and a good job at predicting the mean number of arrests by the end of the follow-up period as shown in Table 7.5.

Before describing the trajectories in Figures 7.2 and 7.3, it is important to point out to our readers that the goals of the description of the offending trajectories are focused on addressing the key issues of this chapter: (1) whether there are offender groups in the population with distinct arrest trajectories; (2) whether there is an "adolescent-limited" group in the serious youthful offender population; and (most importantly) (3) whether between-group differences are maintained across time. It is critical to note at this point that the substantive implications of both the predicted and

[18] The reason the predicted trajectory undercuts four of the actual observed peaks (LC1, LC2, LC3, and LC4) is that the rate of increase in the arrest during early adolescence is much faster than the corresponding decline in adulthood (which is much more drawn out over time) for these four latent classes. Mathematically speaking, such a trend makes it extremely difficult to accurately model the arrest rate at each and every age, and the model favors undercutting the peak rate because there are simply more data points during adulthood (when the decline is slower). On average, though, it is important to remember that the model does an excellent job of predicting the final number of arrest charges

observed arrest trajectories are identical, even if the peak ages are not always identical. For the reader interested in the comparison of the observed and predicted arrest trajectories for each latent class, we direct your attention to Appendix F which contains a graphical depiction of these trajectories.

We begin our discussion of Figures 7.2 and 7.3 by noting that it is readily apparent from the trajectories in both figures that there are groups with very distinct arrest trajectories across the age distribution studied here. The LC4 latent class, for example, had a very low-rate across time, whereas LC6 had a very high-rate. Thus, these offender groups are not only statistically different (as the results presented earlier indicated), but there are also significant substantive differences between these offender groups as well (which are discussed further below).

The substantive nature of the arrest trajectories presented in Figure 7.2 and 7.3 lend support to hypothesis H_1 that there are multiple offender groups in the serious youthful offender population, and these results fail to support hypothesis H_2 predicting that there are only two distinct offender groups. There is simply more heterogeneity (even within this select segment of the offender population) than that which is expected on the basis of the predictions of Moffitt's dual taxonomy theory (1993, 1997).

Second, note that the arrest trajectory of the LC5 latent class clearly follows an adolescent-limited trajectory.[19] Thus, even in this sample of very serious offenders, there is a group of offenders for whom criminal behavior clearly appears to decline over the age curve. This group of offenders (in terms of both their actual and observed rates) had

[19] As noted by D'Unger et al. (1998), a trajectory such as the LC5 trajectory would be more accurately labeled as "adolescent-peaked" than "adolescent limited," but to maintain continuity with the argument of Moffitt, we refer to this trajectory as adolescent-limited.

a rapid increase in their rate of arrest once onset began (which on average was when they were 13.57 years old), and then there was a nearly identical decrease in their arrest rate after their peak age. By their early twenties, this group appears to have largely desisted from offending. In fact, for the individuals assigned to this trajectory, there were only a handful of arrests between the ages of 21 through 26. From ages 27 through 42, there was not even a single arrest charge for the offenders assigned to this latent class. Thus, the (predicted and observed) arrest trajectories of the LC5 latent class clearly lend support to hypothesis H_3 that there is an adolescent-limited offender group in the very serious youthful offender population.

Finally, it is also clear from Figures 7.1 and 7.2 that the trajectories of the offender groups depicted clearly differ not only with respect to their average rate of arrest, but that they are also substantively different in terms of the growth and decline of their arrest trajectories over time. In other words, the between-group differences are not stable over time. By virtue of the fact that there is an adolescent-limited group, this conclusion is relatively straightforward (i.e., notice how this trajectory cuts across all of the other arrest trajectories). But, it is important to note that even were we to exclude the adolescent-limited latent class from our analysis, we would still observe a lack of stable between group differences among the remaining classes. The between-group differences in our sample are largely maintained up though age 15, but after age 15 (and throughout adulthood), there is a clear failure to maintain these between-group differences. For example, note that the LC1 latent class had a much lower and later peak age of arrest in comparison with the LC3 latent class, but at about age 25 the LC1 latent class was predicted (and observed) to have a higher arrest rate than did the LC3 latent class. In

fact, by their mid 30s, the LC1 latent class had a rate of arrest very similar to that of LC2, even though during earlier ages the arrest rate of LC2 was much higher than was the arrest rate of the LC1 latent class. Similarly, notice that the LC6 latent class, which had both the highest rate of arrest during adulthood and the highest average arrest charge total overall (see Table 7.5), fell right in the middle of both the observed and predicted arrest trajectories all the way though age 16. Shortly thereafter, however, this latent class assumed the rank of the one with the highest arrest rate at the later adult ages.

In short, the arrest trajectories depicted in Figures 7.2 and 7.3 reject the notion that between-group differences are maintained across time. Thus, at this point we have seen both substantively and statistically that the general evidence from the analyses reject the notion that the relationship between age and crime is "invariant" across the latent classes.

The Adult Prison Experiences of the Latent Classes

In this final section for the 1981-82 sample, we present some descriptive evidence concerning the adult prison experiences of the latent classes in this sample. The reason for this presentation is that a rival hypothesis for the failure of the maintenance of the between-group differences in recidivism is that some of the groups had distinctively different prison experiences. For example, it could be argued that the adolescent-limited group was less likely to have been arrested from their mid 20s through the end of the follow-up period in 2000 because they were more likely to have been incapacitated in prison and were thus denied the opportunity to offend against the non-institutionalized public. The presentation here is merely descriptive—future analyses will have to be

undertaken to more definitively address the role that adult incarceration time plays in determining the nature of the arrest trajectories.

Table 7.6 presents a brief summary description of the adult incarceration experiences for each of the latent classes in the 1981-82 sample, including the percentage of each latent class that had at least one "stay" in the CDC (California Department of Corrections), the percentage that was incarcerated at the end of the follow-up period, and among those who had at least one stay, the average number of CDC "stays" and the median years spent incarcerated in the CDC.

As depicted in Table 7.6, there is a marked difference in the percentage of each latent class that has at least one incarceration in the CDC. The adolescent-limited group has the lowest percentage (15.81%), whereas the LC2 and LC6 latent classes have the highest percentages (both around 89%). Not surprisingly, the results in Table 7.6 show the stochastic nature of the process through which offenders end up incarcerated in the prison system—the higher the number of criminal arrests (during adulthood), the higher the rate of entrance into prison (see e.g., Canela-Cacho et al. 1997). Further, the mean number of stays in the CDC and the median number of years incarcerated in the CDC obviously also are highly correlated with the number of criminal arrests during adulthood. The latent classes that had the most active arrest records during adulthood (LC1, LC2, LC3, LC6) also tended to have the highest mean number of prison entrances and spent the most amount of time incarcerated.

Again, the results presented in Table 7.6 appear to indicate the obvious in that they demonstrate that those who were most frequently arrested as adults were the ones who were most likely to be sentenced to prison and have the longest prison stays. We

Table 7.6 Summary of Variables Related to Adult Incarceration Experiences, by Latent Class: 1981-82 Sample

Variable	Latent Class					
	LC1	LC2	LC3	LC4	LC5	LC6
% With At Least 1 CDC Stay	73.03	88.83	71.74	21.72	15.81	89.18
% Still Incarcerated on June 30, 2000	13.87	21.34	16.43	5.74	5.58	6.19
Among Those With at Least 1 CDC Stay:						
Avg # of CDC Stays	2.39	3.08	2.02	1.64	1.21	2.85
Median # of Total Years Incarcerated in CDC	4.02	5.26	5.36	2.92	4.25	3.01

suspect, however, that such findings do not offer confirmation for the notion that the causes of the between group differences in arrest rates *over the entire age distribution* were due merely to differential adult incarceration rates and lengths of prison stays. For example, if the rates of arrest decline in adulthood were purely a function of time spent incarcerated in prison, then why do the LC1 and LC3 latent classes have the same percentage that make a transition into prison, and nearly the same median amount of time incarcerated for those that had at least 1 stay (differing only by about 1 year in total out of a possible 18 years of exposure time after release from the CYA), yet vastly different rates of arrest decline in adulthood? If incapacitation explained arrest differentials between the two latent classes, you would expect the two groups to have similar declines in arrest rates since their imprisonment experiences are vary similar. Furthermore, the adolescent peaked group had the lowest rate of prison entrance, yet they had the most dramatic decline in the adult arrest patterns. It is, of course, highly suggestive to contend that adult prison experiences were not driving the differences in the arrest rates between the latent classes over the entire age distribution. Future analyses that are able exclude time incarcerated in prison from their measures of time "at-risk" for arrest will permit a more definitive examination of this issue.

RESULTS FOR THE 1986-87 SAMPLE

In this section, we present the "substantive" results for the 6-class semiparametric mixed Poisson model of the 1986-87 sample. These analyses are procedural replications of the same processes employed in the results section presented above for the 1981-82 sample.

Latent Class Assignment Percentages and Posterior Probabilities

Our analysis of data for the 1986-87 sample begins with a presentation of the percentage of cases assigned to each latent class and a descriptive summary of the posterior probabilities of assignment to the latent class. The latent classes were named according to the same convention used above for the 1981-82 sample. The latent class with the largest number of estimated "members" was labeled LC1, descending down to the latent class that was estimated to have the fewest members belonging to it, which was labeled LC6.

Table 7.7 contains two panels: Panel A presents the estimated and actual group assignment percentages and Panel B contains a summary description of the posterior assignment probabilities. Similar to the results reported for the 1981-82 sample, there are no latent classes that contain an overwhelming majority of the cases. Using the maximum probability assignment rule (i.e., each cases was assigned to the latent class to which they had the highest probability of belonging), the following percentage of cases was assigned to each specific latent class: 28% (LC1; n = 398); 23% (LC2; n = 333); 15% (LC3; n = 211); 13% (LC4; n = 188); 12% (LC5; n = 170); and 10% (LC6; n = 143).

Table 7.7. Summary Descriptions of Latent Class Assignment Percentages and Posterior Assignment Probabilities: 1986-87 Sample

Panel A: Group Assignment Percentages

Latent Class	Estimated %	Assigned %	Assigned N
LC1	27.08	27.58	398
LC2	23.23	23.08	333
LC3	15.09	14.62	211
LC4	12.58	13.03	188
LC5	11.88	11.78	170
LC6	10.13	9.91	143

Panel B: Posterior Assignment Probabilities

Latent Class	Mean	Percentiles 25th	Percentiles 50th	Percentiles 75th
LC1	0.87	0.80	0.95	0.99
LC2	0.86	0.77	0.95	0.99
LC3	0.82	0.66	0.88	0.99
LC4	0.93	0.92	0.99	1.00
LC5	0.85	0.77	0.94	1.00
LC6	0.88	0.82	0.96	1.00

There is also a model-based estimate of the percentage of cases that were expected to belong to each latent class found in Table 7.7. Comparing the estimate percentage to the actual assigned percentage in each latent class, we find a considerable degree of similarity between the two percentages. The largest discrepancy between the estimated and actual percentages was less than 1% (0.50% in LC1).

Panel B of Table 7.7 contains a summary description of the posterior assignment probabilities. Each row of this panel pertains to a description of the assignment probabilities for the individuals who were actually assigned to that latent class. Although not as high as the assignment probabilities found in Table 7.4 for the 1981-82 sample, the average assignment probabilities for the 1986-87 sample were still quite high. Most of the average assignment probabilities ranged between 0.85 and 0.93. The latent class with the lowest average assignment probability was LC3, which had an average assignment probability of 0.82.

Looking at the medians (50th percentiles) presented in Panel B of Table 7.7, we find that the median values for most latent classes were very high. The lowest median probability value for any of the latent classes was 0.88, which indicates that over 50% of the individuals in the 1986-87 sample were assigned to the latent class that had at least a 0.88 probability of having generated the individual's longitudinal offending pattern.

Summary Arrest Charge Information

In this section we present a descriptive summary of the arrest histories of the individuals in the 1986-87 sample assigned to each latent class. Table 7.8 presents the observed and predicted mean number of arrest charges, the mean number of observed

Table 7.8 Summary Arrest Charge Information, by Latent Class: 1986-87 Sample

Latent Class	Mean Total Arrest Charges		Obs. Mean Serious Arrest Charges	Average Age at First Arrest
	Observed	Predicted		
LC4	9.45	9.46	5.37	14.27
LC1	14.55	14.59	8.52	14.03
LC2	18.31	18.41	9.91	14.63
LC3	27.62	27.19	14.99	11.29
LC6	35.87	35.46	16.61	13.59
LC5	36.19	35.75	17.95	13.40

serious arrest charges, and the average age at first arrest for each of the latent classes. As discussed earlier in Chapter 6, the "average individual" in the 1986-87 sample was arrested for 21.33 total charges, of which 11.29 were for serious offenses. The average individual in this sample was first arrested when he was 13.7 years old. Similar to the results presented for the 1981-82 sample, we report the information in Table 7.8 rank-ordered by the mean observed number of arrest charges (from lowest to highest).

As expected, given the findings presented above for the 1981-82 sample, there is a considerable amount of heterogeneity in the average number of total arrest charges accumulated in each latent class for the 1986-87 sample as well. The mean number of total arrest charges varied from a low of 9.46 in the LC4 latent class (which accounted for 13% of the sample), while the highest average number of arrest charges was 36.19 in the fifth latent class (LC5; 12% of sample). The sixth latent class, LC6 (10%), was not far behind the LC5 latent class in terms of their average arrest charge total. The LC6 latent class averaged 35.87 total arrest charges. The difference in the mean number of arrest charges between the latent classes with the highest and lowest arrest charge averages was nearly 27 arrest charges. The LC1 latent class with the largest percentage of the sample assigned to it (27%) had an average of 14.55 arrest charges. This is well below the overall average number of arrest charges for the sample as a whole (21.33). In the remaining two latent classes, the average arrest charge totals were 18.31 (LC2; 13%) and 27.62 (LC2; 15%), respectively. Similar to the 1981-82 sample, we find that the average arrest charge totals within each latent class for the 1986-87 sample are quite a bit different when compared to the average number of arrest charges first presented in Chapter 6 for this sample as a whole.

The "Predicted" charge column in Table 7.8 contains the model-based predicted number of total arrest charges based on the estimates from the semiparametric mixed Poisson model. Comparing this average total to the observed average total for each latent class, we find that the model was fairly accurate, on average, in predicting the sum number of arrest charges in each of the latent classes. For example, the model predicted the LC4 latent class would have 9.46 arrest charges, and the individuals assigned to this latent class were observed to have on average, 9.46 total arrest charges. Similarly, for the latent class with the highest observed number of arrest charges (LC5), the model predicted they would average 35.75 arrest charges and they were found to have on average, 36.19 arrest charges. Overall then, the model does a fairly accurate job at predicting the total number of arrest charges by the end of the follow-up period in the 1986-87 sample. This is the same conclusion we reached for the 1981-82 sample.

Again, as in the 1981-82 sample, serious arrest charges comprised the majority of arrest charges for most of the latent classes in the 1986-87 sample. The latent class with the lowest percentage of serious charges was the LC6 latent class. The LC6 latent class was arrested, on average, for 35.87 charges, and this latent class averaged 16.61 serious charges. In total, then, 46% of this latent class' average arrest charges were for serious offenses. The latent class with the highest percentage of serious charges was the LC1 latent class (59%). Overall though, in each latent class roughly one-half of the arrest charges were for felonies.

Just as we have found heterogeneity between the latent classes with regard to the average total number of arrest charges, we also find heterogeneity between the latent classes in terms of their average age at first arrest. The LC3 latent class had the youngest

average age at first arrest—they averaged their first arrest at the precocious age of 11.29 years old. The LC6 and LC5 latent classes both averaged their first arrest at around 13.50 years old, while the LC1, LC2, and LC4 latent classes were, on average, about 14 years old when they were first arrested.

Similar to the findings for the 1981-82 sample, there is no orderly reverse-ordering of the average ages at first arrest and the rank-orderings (from lowest to highest) of the total arrest charges for the 1986-87 sample. For example, note that the latent class with the youngest average age at first arrest (LC3) did not accumulate the highest average number of arrest charges. Similarly, the latent class with the oldest average age at first arrest also did not accumulate the lowest number of average arrest charge totals. Indeed the two latent classes with highest average arrest charge totals (LC5 and LC6) both had average ages at first arrest that were nearly identical to the average age of arrest for the sample as a whole (13.68 years old). Again, these results are consistent with those presented earlier that tested (and rejected) the hypothesis that the estimated age parameters were equal across the latent classes in this sample (Table 7.3), and they are inconsistent with the "age invariance hypothesis" promulgated by Gottfredson and Hirschi (1990). To more definitively investigate the substantive differences in the arrest trajectories of these latent classes, we present in the next section a graphical depiction of the observed and predicted arrest trajectory for each latent class.

Predicted and Observed Trajectories of Arrest for the 1986-87 Sample

Similar to the analysis of data for the 1981-82 sample, we present two figures that help describe the arrest trajectories for the 1986-87 sample. Figure 7.4 contains the

Figure 7.4. Predicted Arrest Trajecories: 6-Class Model of the 1986-87 Sample

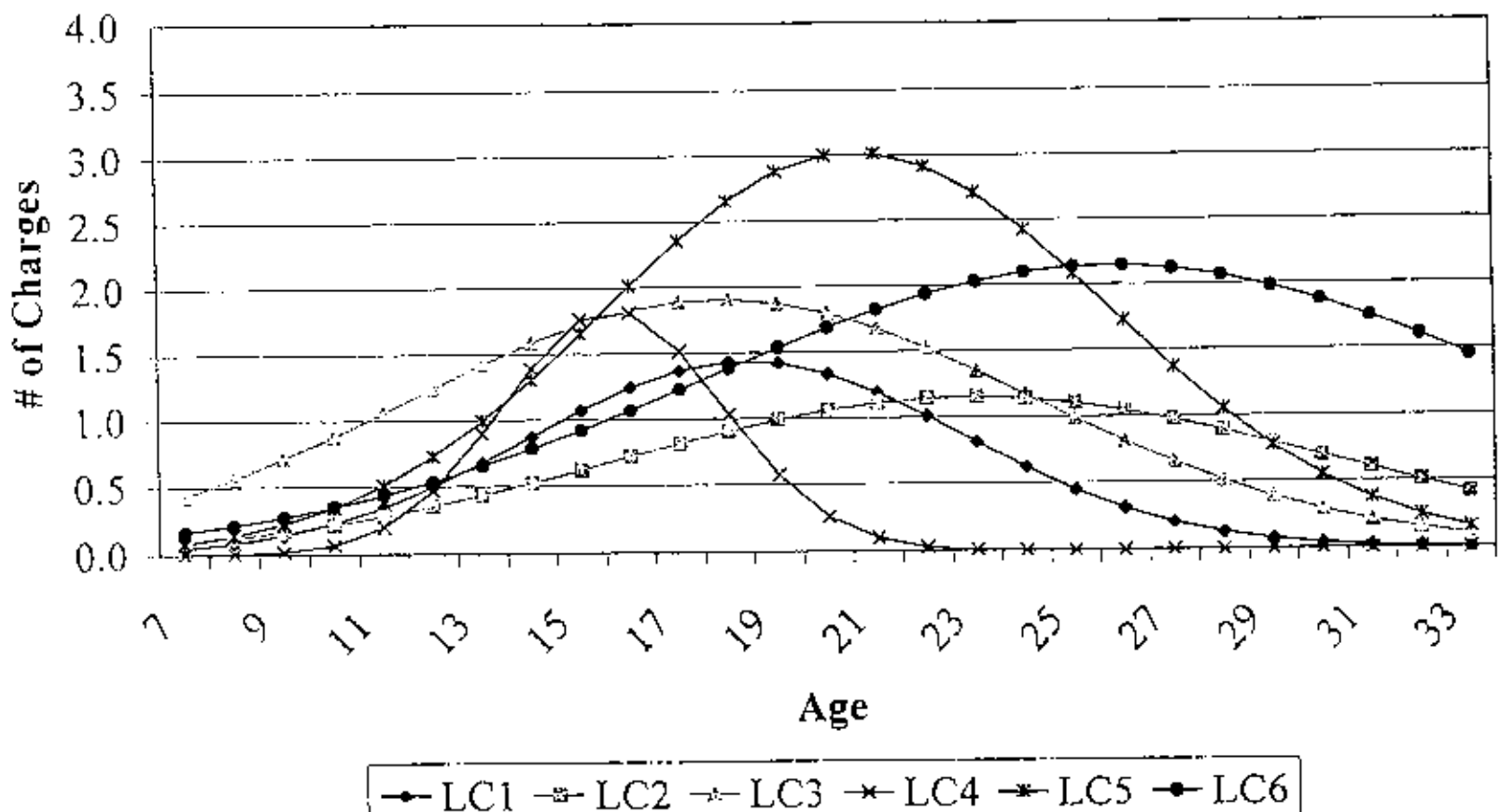
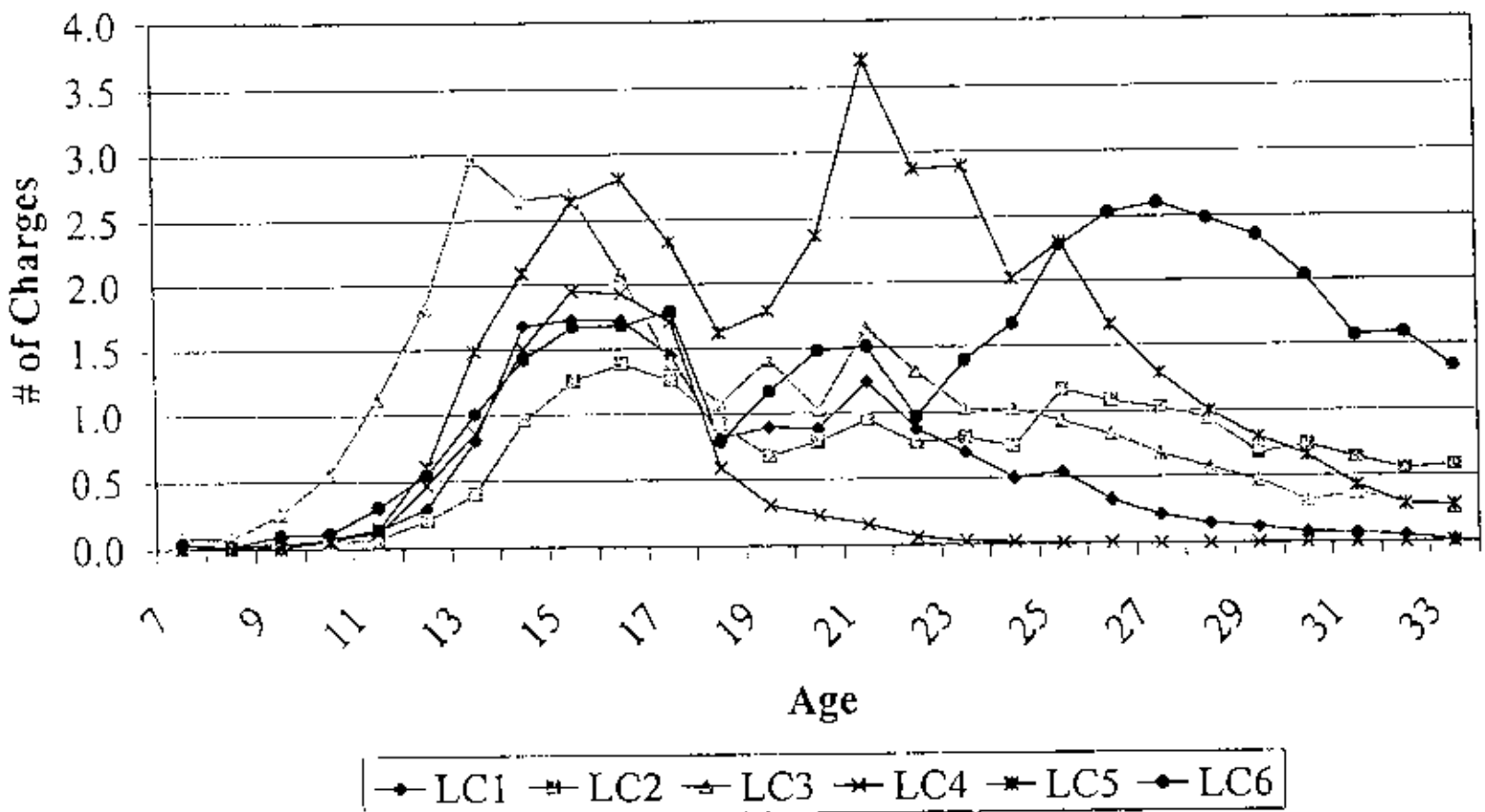


Figure 7.5. Observed Arrest Trajectories: 6-Class Model of the 1986-87 Sample



predicted trajectories and Figure 7.5 the actual observed trajectories for the six latent classes in the 1986-87 sample. The observed arrest trajectories presented in Figure 7.5 were generated by averaging the arrest charge totals at each age for all of the individuals in a given latent class that were at risk at each age.

Similar to the results presented for the 1981-82 sample, the predicted trajectories here also tend to slightly over-estimate the arrest rates at the younger ages, undercut the peak ages of arrest, and slightly over-predict arrests during the early 20s. Again, the most probable reason for the dip in arrests *during the early twenties* is that it is appears to be a "time at risk" artifact due to an unmeasured drop in the "at risk" time as a result of increased likelihood of incarceration (for both new offenses and parole revocations) during this age period (see Footnote #4 of this chapter). It is important to note, that whatever the cause of the drop, the observed number of charges are back "on track" with the predicted number of charges by the time the 1986-87 sample reaches its mid-twenties.

As in the description of the results for the 1981-82 sample, we focus the discussion of the trajectories represented in Figures 7.4 and 7.5 around the key issues of the age-crime relationship identified earlier. For the reader interested in the comparison of the observed and predicted offending trajectories for each latent class, we refer you to the graphical depiction in Appendix F. Due to the fact that it is easier to see the differences in the predicted trajectories (because of the natural "smoothing" that occurs in the predictions), we focus our discussion on these trajectories that are presented in Figure 7.4. The substantive implications of both the predicted and observed trajectories are identical, and the fact that the peak ages of arrest are not always identical in the two figures is immaterial to the main focus of our study.

First, in both of the figures it is clear (although it is clearer in Figure 7.5) that the latent classes have qualitatively distinct arrest trajectories. For example, the LC5 latent class had a very high peak rate of offending, whereas the LC2 latent class had a much lower peak rate. Again, we find that the nature of the trajectories presented in these figures lends substantive support to the first hypothesis (H_1) that there are multiple offender groups in the serious youthful offender population. Also, these results appear to contradict the second hypothesis (H_2) that there are only two distinct offender groups in this population. It would be very hard to envision that all of these different arrest trajectories could be adequately described using only two trajectories (which would be one step beyond Figure 7.1). There is simply too much heterogeneity between these latent classes in terms of both their mean rate of offending and the developmental nature of their arrest trajectories to fully capture the heterogeneity with only a mere two latent classes.[20]

Second, we also find once again that there is an adolescent-limited offender group (LC4) in the serious youthful offender population. The trajectory of the LC4 latent class clearly follows an adolescent-limited trajectory, with a rapid increase in their arrest rate during early adolescence, and then a nearly identical decrease in the arrest rate on the other side of their peak age (age 16). By their early twenties, this group (like the LC5 latent class in the 1981-82 sample) had largely desisted from offending. For the individuals assigned to this trajectory, there were only a handful of arrests between the ages of 21 through 26. From ages 27 through 37, there was not even a single arrest charge

[20] For example, the adolescent-limited latent class does not even get "extracted" until a 5-class model. Thus, without allowing for more heterogeneity in the population than simply two classes, the fact that there is an adolescent-limited trajectory would be completely lost.

for the offenders assigned to this latent class (that were followed through those ages). Thus, the trajectory of the LC4 latent class lends support to third hypothesis (H_3) that there is an adolescent-limited offender group in the serious youthful offender population.

Finally, the trajectories depicted in both Figures 7.4 and 7.5 provide more evidence against the "age invariance" hypothesis put forth by Gottfredson and Hirschi. Whether you look at the observed or the predicted arrest trajectories, it is clear that the trajectories depicted in these figures do not maintain their differences over time. Parallel to the findings discussed above for the 1981-82 sample (and excluding the adolescent-limited group), there was more stability in the differences among the trajectories during the early ages studied here (through age 15). There was, however, a significant change in the between-group differences in arrests throughout adulthood. For example, note that the LC3 latent class had the highest arrest rate during the early ages studied here (through about age 15), but then by the late 20s there were four latent classes with higher predicted (and observed) arrest trajectories. By itself, the mere presence of an adolescent-limited trajectory poses a serious problem to the age-invariance hypothesis. The arrest trajectory of the adolescent-limited group is simply incompatible with the age-invariance argument because this arrest trajectory of the adolescent-limited offender group (LC4) drops right across the (predicted and observed) arrest trajectories of several of the other arrest trajectories (e.g., LC1, LC2, LC6). In other words, their arrest rate was at one point significantly higher than the arrest rate of the other latent classes, but by the early twenties the LC4 trajectory is the only trajectory that is running along the X axis at a predicted (and observed) rate of zero. Or consider the arrest trajectory for the LC2 latent class. This latent class had the lowest arrest rate through age 19, but by age 33 it had the

second highest arrest rate. Finally, another arrest trajectory that is problematic for the age invariance hypothesis is that of the LC6 latent class. The LC6 latent class displayed an offending trajectory that was midway between the other arrest trajectories through about age 17 (both in terms of the observed and predicted trajectory). By the end of the follow up period, this latent class had the highest arrest trajectory (both predicted and observed).

Overall, the trajectories depicted in Figures 7.4 and 7.5 reject the notion that between-group differences in arrest are maintained across time. As in the 1981-82 sample, the results to this point in our analyses have provided both substantive and statistical evidence that reject the notion that the relationship between age and crime is "invariant" across the latent classes.

The Adult Prison Experiences of the Latent Classes

In this final section of results for the 1986-87 sample, we again present some descriptive evidence concerning the adult prison experiences of the latent classes, this time with the second sample. Again, the purpose of this presentation is to examine descriptively whether the between-group differences appear to be the result of differential incarceration experiences. Table 7.9 includes the percentage of each latent class that had at least one "stay" in the CDC, the percentage that was incarcerated at the end of the follow-up period and those who had at least one stay, the average number of CDC "stays" and the median years spent incarcerated in the CDC.

Like the results observed for the 1981-82 sample, Table 7.9 depicts a striking difference in the percentage of each latent class that had at least one entrance into the CDC. The percentages range from 14.4% (LC4) all the way up to 92.2% (LC5). The

Table 7.9 Summary of Variables Related to Adult Incarceration Experiences, by Latent Class: 1986-87 Sample

Variable	Latent Class					
	LC1	LC2	LC3	LC4	LC5	LC6
% With At Least 1 CDC Stay	64.14	78.62	84.88	14.44	92.22	80.43
% Still Incarcerated on June 30, 2000	11.62	9.75	13.66	5.35	17.37	6.52
Among Those With at Least 1 CDC Stay:						
Avg. # of CDC Stays	1.74	2.07	2.29	1.15	2.6	2.29
Median # of Total Years Incarcerated in CDC	4.01	2.71	4.19	4.35	3.48	2.06

mean number of stays in the CDC (among those who had at least one stay) varied from 1.15 (LC4) to 2.6 (LC5), while the median number of total years incarcerated (among those incarcerated at least once) ranged for 2.06 (LC6) to 4.35 (LC4).

For our purposes here, however, again we note that the results presented Table 7.9 do not support the possible interpretation that the differences were most likely the result of possible incapacitation effects *over the entire age distribution*. For example, again we find that the adolescent-limited offender group (which experienced the most abrupt change) had an extremely low-rate of entrance into the CDC (which makes sense given their lack of arrests during the majority of the adult years). Again we note that LC5 and LC6 both had extremely high rates of entrance into the CDC, yet their arrest trajectories did not respond in identical manners.

To look at this finding a different way, the reader can examine the percentage of cases that make an entrance for the LC1 and LC3 latent classes. Sixty-four percent of the LC1 latent class had at least one stay in the CDC (and those sixty-four percent spent 4 years in the CDC according to the median), whereas 85% of the LC3 latent class makes a transition into the CDC (and spends a median length of 4 years there). Thus, 20% more of the LC3 latent class made a transition into the CDC, but the decline during adulthood for these two arrest trajectories was virtually identical. If incapacitation effects were causing the arrest trajectories to fall at different rates, then the LC3 trajectory should have fallen at a much faster rate than the arrest trajectory of the LC1 latent class. Stated differently, why do the trajectories of these two latent classes change at the same rate when they had different CDC experiences (20% more of the LC3 latent class served time compared to LC1 group)?

Whether you consider why the group with the most significant and fastest drop in arrests has the lowest prison entrance rates, or why two groups with the same change in arrests have different CDC experiences, the conclusion that differential incarceration rates were causing the between group differences to be reduced does not make sense, given the descriptive results depicted in Table 7.9. Again, these results merely suggest that differential incarceration rates were not causing the different changes in the shapes of the arrest trajectories, but a more thorough (and definitive) examination of this issue is warranted in future research.

RESULTS FOR THE 1991-92 SAMPLE

This final section presents the "substantive" results of the 6-class semiparametric mixed Poisson model for the 1991-92 sample. The analyses that follow here are identical to those performed above for the previous two samples.

Latent Class Assignment Percentages and Posterior Probabilities

The analysis of the data for the 1991-92 sample begins with a presentation of the percentage of cases assigned to each latent class and a descriptive summary of the posterior probabilities of assignment to the latent class. The latent classes naming conventions are identical to those used above.

Panel A of Table 7.10 presents the estimated and actual group assignment percentages, and Panel B summarizes the posterior assignment probabilities for the 1991-92 sample. Using the maximum probability assignment rule, the following percentages

Table 7.10. Summary Descriptions of Latent Class Assignment Percentages and Posterior Assignment Probabilities: 1991-92 Sample

Panel A: Group Assignment Percentages

		Assigned	
Latent Class	Estimated %	%	N
LC1	27.67	27.62	396
LC2	23.2	23.99	344
LC3	15.85	15.62	224
LC4	14.78	14.71	211
LC5	10.65	11.09	159
LC6	7.85	6.97	100

Panel B: Posterior Assignment Probabilities

		Percentiles		
Latent Class	Mean	25th	50th	75th
LC1	0.84	0.74	0.91	0.98
LC2	0.82	0.69	0.87	0.96
LC3	0.86	0.79	0.94	0.99
LC4	0.82	0.65	0.91	0.99
LC5	0.89	0.79	0.96	1.00
LC6	0.87	0.76	0.96	1.00

of cases were assigned to each latent class: 28% (LC1; n = 396); 24% (LC2; n = 344); 16% (LC3; n = 224); 15% (LC4; n = 211); 11% (LC5; n = 159); and 7% (LC6; n = 100).

The model-based estimate of the percentage of cases that were expected to belong to each latent class can also be found in Table 7.10. Looking at both the estimated and actual assigned percentages, we find a remarkable degree of similarity between the two percentages. Just like the two previous samples, the largest discrepancy between the two percentages was less than 1% (0.79% in LC2).

The summary description of the posterior assignment probabilities can be found in Panel B of Table 7.10. Again, each row of this panel describes the posterior probabilities of only those individuals who were actually assigned to that latent class. The average assignment probabilities for all of the latent classes were between 0.82 and 0.89. Given that there are significantly fewer "age years" or "data points" included in the analytical data file of the 1991-92 sample compared to the analytical data files of the previous samples, it is not surprising that the posterior assignment probabilities are not as high in this sample (i.e., the more "trials" in the data, the more information there is to compute the posterior probabilities). The two latent classes with the lowest average assignment probabilities for the 1991-92 sample were the LC2 and LC4 latent classes, both of which had average assignment probabilities of 0.82. An examination of the medians (50th percentiles) presented in Panel B of Table 7.10 indicates that most latent classes had median assignment probabilities that were fairly high—the lowest median probability value was only 0.87. Thus, over 50% of the individuals in the 1991-92 sample were assigned to the latent class that had a 0.87 probability or greater of having generated the individual's longitudinal offending pattern.

Summary Arrest Charge Information

At the outset, we note here that many of the substantive findings of this section concerning the overall nature of the arrest patterns of the latent classes are, in fact, virtually identical to the findings noted above in both the 1981-82 and 1986-87 samples. Therefore, some of the findings are not discussed in as much detail as those presented previously.

A descriptive summary of the arrest charge histories of the individuals assigned to each latent class for the 1991-92 sample can be found in Table 7.11. The information presented in Table 7.11 includes the average observed number of total arrest charges (at the end of the follow-up period on June 30, 2000), the predicted average number of arrest charges based on the estimates from the semiparametric mixed Poisson model, the observed number of serious arrest charges, and the average age at first arrest. To briefly review our previous description of the overall arrest patterns for the sample as a whole (from Chapter 6), the "average individual" in the 1991-92 sample accumulated 16.56 arrest charges by the end of the follow-up period, of which 8.59 of those were for serious arrest charges. The "average" individual was first arrested at an age of 13.64 years old. As in the two previous sample results sections, the information depicted in descriptive arrest history (Table 7.11) is rank-ordered by the mean number of observed arrest charges (from lowest to highest) because this makes for a clearer presentation of some of the other information in the table.

The results displayed in Table 7.11 speak to a significant amount of between-group (or between-class) heterogeneity in the average number of arrest charges. This finding was expected given the earlier findings obtained for both the 1981-82 and 1986-

Table 7.11. Summary Arrest Charge Information, by Latent Class: 1991-92 Sample

Latent Class	Mean Total Arrest Charges		Obs. Mean Serious Arrest Charges	Average Age at First Arrest
	Observed	Predicted		
LC5	9.43	9.63	5.50	13.88
LC2	11.36	11.42	6.37	14.54
LC1	13.06	13.03	7.23	14.22
LC4	19.12	18.78	10.20	11.35
LC6	27.26	26.53	12.87	13.90
LC3	28.57	28.10	13.18	13.13

87 samples. The mean number of total observed arrest charges varied from a low of 9.43 in the LC5 latent class (which accounted the 11% of the sample), while the highest average number of arrest charges was 28.57 in the fifth latent class (LC3; 16% of sample). As in the 1986-87 sample, there were two latent classes with very high arrest charge totals—the sixth latent class, LC6 (7%), was not far behind the LC3 latent class in terms of their average arrest charge total (27.26). Overall though, there was a difference of over 19 arrest charges between the latent classes with the highest and lowest arrest charge averages. The LC1 latent class with the largest percentage of the sample assigned to it (28%) had an average of 13.06 arrest charges. The average arrest charge totals in the remaining two latent classes were 11.36 (LC2; 24%) and 19.12 (LC4; 15%). As in the two previous samples, the variable average arrest charge totals within each latent class indicate that the average number of arrest charges calculated for the sample as a whole is not very representative of the average number of arrest charges found in each of the distinct latent classes.

The "Predicted" column in Table 7.11 contains the model-based predicted number of total arrest charges based on the estimates from the semiparametric mixed Poisson model. Comparing the averages in the two columns, we find again that the model-based prediction was fairly accurate in predicting the overall average number of arrest charges in each of the latent classes. As in the earlier samples, about 50% of the arrest charges in each latent class were composed of serious arrest charges (the percentages ranged from 46% to 58%).

Once again we find between-class heterogeneity with respect to the average age at first arrest (Table 7.11). The youngest average age at first arrest was found in the LC4

latent class—they averaged their first arrest at the early age of 11.35 years old. The LC3, LC5, and LC6 latent classes averaged their first arrest when they were 13 years old, while the LC1 and LC2 latent classes were about 14 years old when they were first arrested by law enforcement authorities.

Also consistent with the findings from the two previous samples, Table 7.11 fails to depict a neat, reverse-ordering of the average ages at first arrest on the basis of the rank-orderings (from lowest to highest) of the total arrest charges among the latent classes. The latent class with the youngest average age at first arrest (LC4) did not accumulate the highest average number of arrest charges, while the latent class with the oldest average age at first arrest also did not accumulate the lowest number of average arrest charge totals. Once again we find that the two latent classes with highest average arrest charge totals (LC3 and LC6) both had average ages at first arrest that were nearly identical to the average age of arrest for the sample as a whole (which is 13.64 years old).

These results support the previously stated conclusion that the estimated age parameters were not equal across the latent classes in this sample (Table 7.3), and they are also inconsistent with the "invariance hypothesis" promulgated by Gottfredson and Hirschi (1990). Evidence supporting the age invariance hypothesis requires that the latent class with the highest average number of arrest charges also have the youngest age of arrest, and similarly that the latent class with the oldest average age at first arrest should have the lowest average number of arrest charges. When presumed invariants such as these display variance across latent classes, it is a clear indication of changing between-group differences over time. In the next section we present a graphical

depiction of the observed and predicted arrest trajectory for each latent class to better investigate the substantive differences in the arrest trajectories of these latent classes.

Predicted and Observed Trajectories of Arrest

As in the prior section, the findings here replicate those noted above in the 1981-82 and 1986-87 samples. As before, two figures are described in this section. Figure 7.6 contains the predicted arrest trajectories, while Figure 7.7 contains the actual observed trajectories for the six latent classes in the 1991-92 sample.

Similar to the findings observed for the two prior samples, the predicted trajectories for the latent classes in the 1991-92 sample here also tend to slightly over-estimate the arrest rates at the younger ages, undercut the peak ages, and slightly over-predict during the early 20s. And once again, regardless of the causes of the drop in arrests, the observed number of charges are back "on track" with the predicted number of charges once the sample members reach their mid-twenties. Again, for the reader interested in the comparison of the observed and predicted arrest trajectories for each latent class, there is graphical depiction of the observed and predicted trajectories for each latent class in Appendix F. Due to the fact that it is easier to see the differences in the predicted arrest trajectories (because of the natural "smoothing" that occurs in the predictions), we focus our discussion on the predicted trajectories in Figure 7.6. However, it is important to note that the substantive implications relevant to this study do not depend on whether one uses the predicted or observed arrest trajectories.

First, both figures 7.6 and 7.7 indicate that the latent classes have qualitatively distinct arrest trajectories (although this depiction is clearer in Figure 7.6). For example,

Figure 7.6. Predicted Arrest Trajectories: 6-Class Model of the 1991-92 Sample

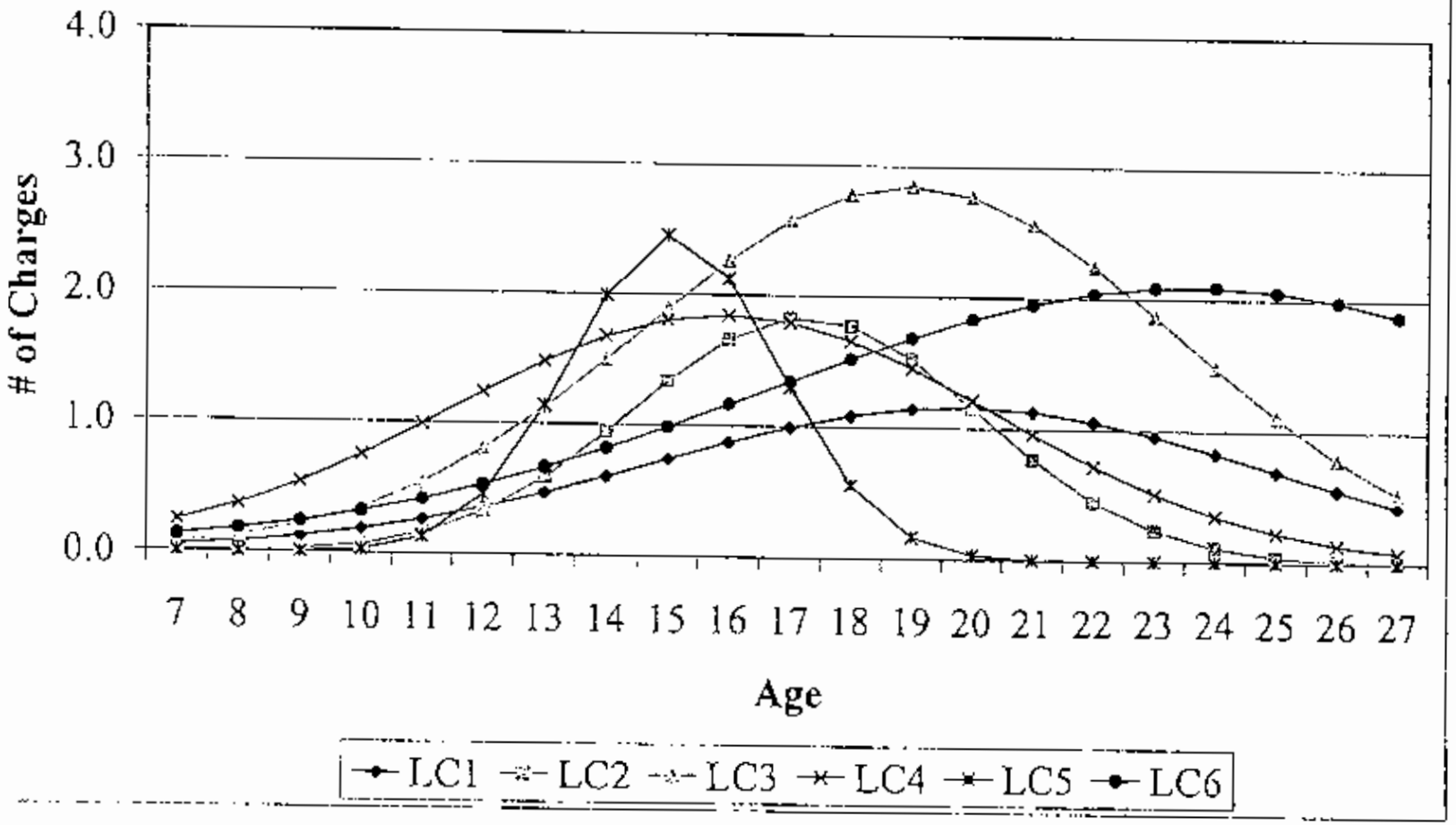
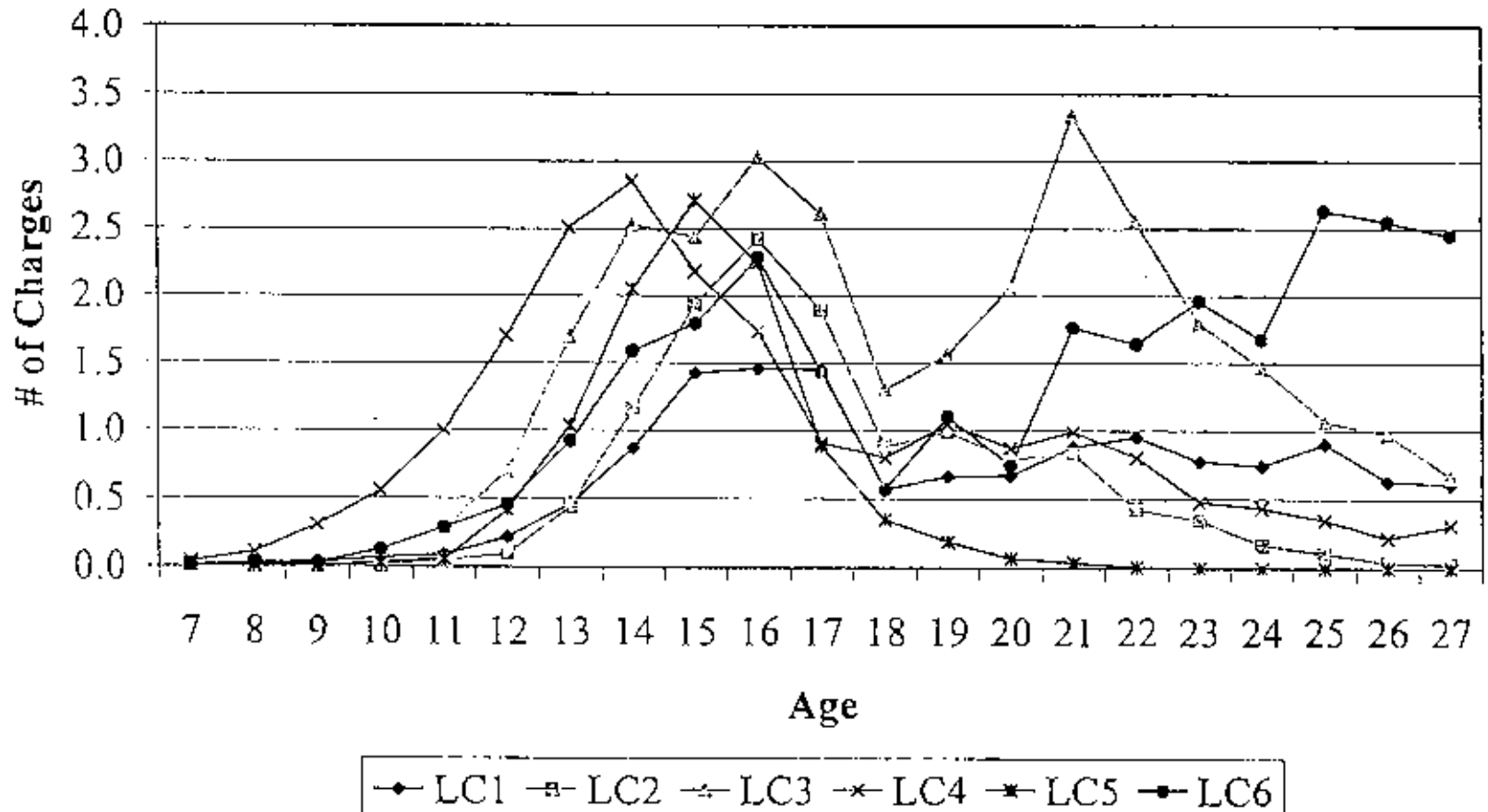


Figure 7.7. Observed Arrest Trajectories: 6-Class Model of the 1991-92 Sample



the LC3 latent class has a very high peak rate of offending, whereas the LC2 latent class has a peak rate that is roughly one-half of the peak rate of the LC3 latent class. The nature of the trajectories presented in these figures lend further substantive support to the first hypothesis (H_1) that there are multiple offender groups in the serious youthful offender population. However, these results also contradict the second hypothesis examined in this study (H_2). Our analyses show that there are more than two distinct offender groups. The heterogeneity among these different arrest trajectories could not be adequately described using only two trajectories. Important differences in the developmental nature of the offending trajectories would be lost if we limited even this select portion of the offender population to contain only two distinct latent classes.[21]

Second, once again we find an adolescent-limited offender group (LC4) in the serious youthful offender population. Thus, the adolescent-limited offender group was uncovered in all three of the samples used in this study. The trajectory of the LC4 latent class clearly follows an adolescent-limited trajectory; there is a rapid increase in their arrest rate during early adolescence, and a nearly identical decrease in the arrest rate on the other side of their peak age (age 16). By their early twenties, this group (like the adolescent-limited groups in the two earlier samples) had largely desisted from offending. There were only a handful of arrests between ages 20 through 22 for the individuals in the LC4 latent class, and from ages 23 through 31, there was not even a single arrest charge for the offenders assigned to this latent class (that were followed through those ages). Thus, once again we find support for the third hypothesis (H_3) regarding the

[21] In this sample as well, the adolescent-limited latent class did not even get added as an offender group in the population until a 5-class model was estimated. Thus, without allowing for more heterogeneity in the population than simply two classes, the fact that there was an adolescent-limited trajectory would be completely hidden

presence of an adolescent-limited offender group in the serious youthful offender population.

Finally, the trajectories depicted in both Figures 7.6 and 7.7 provide further evidence against the "age invariance" hypothesis of Gottfredson and Hirschi (1990). Whether using Figure 7.6 (predicted trajectories) or Figure 7.7 (observed trajectories), there are visible differences in the trajectories of the latent classes in terms of not only their mean rate of offending, but also in the developmental nature of their trajectories as well. Once again we find that (excluding the adolescent-limited group) there was considerable stability in the differences among the trajectories during the early ages (through about age 15). However, there was a significant breakdown in the preservation of between-group differences all the way through the adult years studied here. For example, the LC4 latent class had the highest arrest rate during the early ages, but by the late 20s, three of the other latent class had higher predicted (and observed) arrest trajectories. Further, the presence of the adolescent-limited offenders in the data (LC5) poses a serious problem to the age-invariance hypothesis because the arrest trajectory of the adolescent-limited offender group plunges right across all of the other (observed and predicted) arrest trajectories in the 1991-92 sample. In other words, their rate of arrest was at one point, significantly higher than the arrest rate of the other latent classes, but by the early twenties the trajectory of the adolescent-limited offender group possess the trajectory with the lowest predicted and observed arrest rate (at about zero). The LC6 latent class trajectory is also inconsistent with the age invariance hypothesis. During the early years, this latent class had a trajectory that was in the middle of the other

trajectories, however by the mid-twenties this latent class had a trajectory that was the highest of all the latent classes.

Overall, the trajectory patterns shown in Figures 7.6 and 7.7 reject the notion that between-group differences are maintained across time. As in the two previous samples, both the statistical and substantive results provide evidence to reject the hypothesis that the relationship between age and crime is "invariant" across the latent classes.

The Adult Prison Experiences of the Latent Classes

Finally, Table 7.12 presents a brief summary description of the adult incarceration experiences for each of the latent classes in the 1991-92 sample. Once again, we find that there were varying levels of entrance into the CDC across the six latent classes. Table 7.12 reports that 9% of the adolescent-limited offender group (LC5) made at least one transition into the CDC, whereas 81% of the LC3 latent class made at least entrance into the CDC over the age-period studied here. Among those who made an entrance, the average number of stays ranged from 1.07 (LC5) to 1.79 (LC6), while the median number of years incarcerated in the CDC ranged from 1.60 (LC6) to 2.53 (LC4).

The results presented in Table 7.12 lead to the same conclusion arrived at in the two previous samples. There is no clear (or robust) indication that incarceration differences (or incapacitation effects) *over the entire age distribution* are driving the failure of the groups to maintain their arrest differences over time (again, see Footnote #4). The group that poses the most problems to the "age invariance" hypothesis of Gottfredson and Hirschi (because of their rapid, early decline compared to the other trajectories) is the group with the smallest percentage of its members that were eventually

Table 7.12 Summary of Variables Related to Adult Incarceration Experiences, by Latent Class: 1991-92 Sample

	Latent Class					
Variable	LC1	LC2	LC3	LC4	LC5	LC6
% With At Least 1 CDC Stay	51.40	46.78	80.72	59.72	9.43	62.24
% Still Incarcerated on June 30, 2000	7.38	10.23	9.42	11.37	1.26	6.12
Among Those With at Least 1 CDC Stay:						
Avg. # of CDC Stays	1.57	1.41	1.64	1.54	1.07	1.79
Median # of Total Years Incarcerated in CDC	1.90	2.29	2.07	2.53	2.44	1.60

incarcerated. The other groups, which had much slower rates of change (the kind of change that would be expected when the members were not "incapacitated"), were the groups who had the highest prison entrance rates. Again, future analyses are needed (and will be undertaken) to examine the role played by adult prison stays in determining arrest trajectories more definitively (in a more methodologically robust manner that accounts for time at risk through "offset" terms in the equations). The results here do not indicate that incapacitation effects are driving the loss of between group differences over time.

It is important to note that our conclusion here dovetails with the results of the Piquero et al. (2001) study that examined the adult arrest patterns of a sample of 272 CYA wards and employed the use of "offset terms" in the finite mixture model to account for differences in street time. As Piquero et al. (2001: 68) noted, "the general shape of the arrest trend appears to be robust to controls for exposure time."

COMPARISON OF LATENT CLASSES ACROSS THE SAMPLES

In our review of the prior research on this topic, we noted that, with the exception of the D'Unger et al. (1998) study, most studies of the heterogeneity in longitudinal patterns of criminal activity employing the use of the finite mixture models have generally only had access to a single sample. Thus far, only D'Unger et al. have been able to generalize to the same population over time, and no previous study has yet addressed how "stable" or "unstable" a given set of latent classes are across time. Thus, without more than one dataset, it is impossible to (1) replicate the existence of a trajectory group over time and (2) address whether the nature of the offending trajectories change over time. As such D'Unger et al. (1998) argue that replication of offending

trajectories is a critical research need that is necessary to prevent reifying any particular identified offending trajectory as a stable element in a population.

In fact, the question of whether there are latent classes that are stable elements in a population and fail to change over time is directly related to the question at-hand here. One of the main reasons why Gottfredson and Hirschi's hypothesis (that the relationship between age and crime is invariant across person, place, culture, and time) was such an unpalatable hypothesis to many sociologists (and especially life course sociologists) is that if true, it would imply that social and historical conditions have no formidable impact on crime trajectories. As Benson (2002: 77) notes, "it [invariant age-crime relationship] would call into question the life course principle of contextualism, that is, the idea that social and historical conditions shape trajectories in all domains of life." To life course sociologists, the social context is a "force in development" (Elder and O'Rand 1995) that has the power to redirect or change trajectories already in motion—the long-term shape of a trajectory is by no means fixed to take a particular course after a given age. The age-invariance hypothesis of Gottfredson and Hirschi, on the other hand, envisions a pattern of development that is fixed at a relatively early age. Thus, finding evidence of stable, unchanging latent classes in the population would call into question the idea that changing social conditions are relevant to the developmental shape of a trajectory.

To this point, the focus of this chapter has been on examining and comparing the latent classes discovered within each sample over time. Next, we briefly compare the trajectories across the samples to examine if there were any stable offender groups present in the samples.

Figure 7.8. Graphs of Comparable Latent Class Trajectories Across the Three Samples

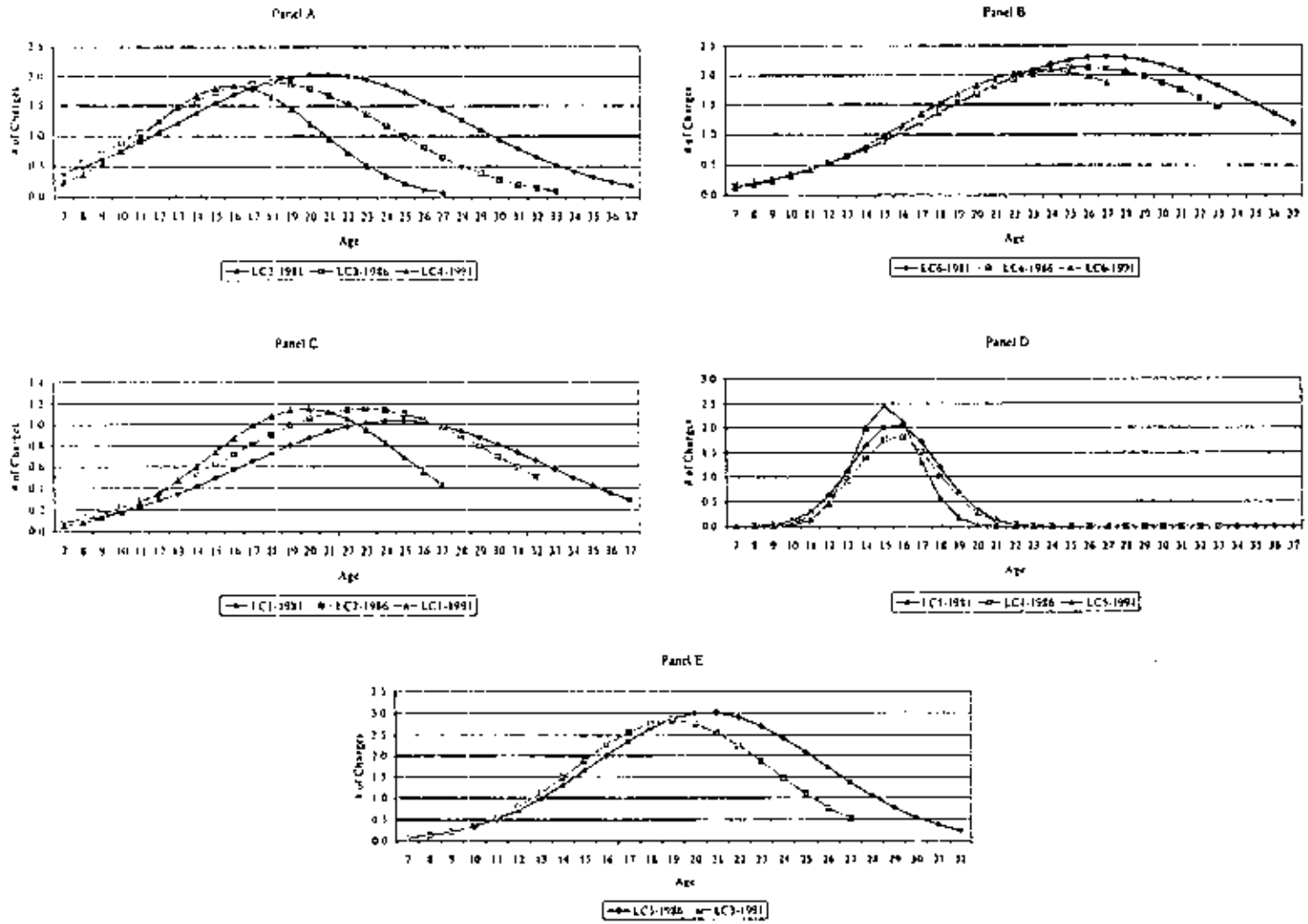


Figure 7.8 contains five panels of predicted arrest trajectories from the results presented earlier. Each panel contains a graphical depiction of the trajectories from each of the latent classes that had "comparable" developmental features.[22] Our examination of the arrest trajectories presented for the three samples indicated that there were four arrest trajectories common to all three samples (Panels A – D of Figure 7.8), and one trajectory that was common to the both the 1986-87 and 1991-92 samples (Panel E of Figure 7.8). Panel A, for example, contains the LC2 trajectory from the 1981-82 sample, the LC3 trajectory from the 1986-87 sample, and the LC4 trajectory from the 1991-92 sample. As depicted in that panel, those three predicted arrest trajectories were nearly identical all the way through age 17, at which point the trajectories began to take different paths. At age 17, the LC4 trajectory of the 1991-92 sample began a much more rapid descent. The LC3 trajectory of the 1986-87 sample followed suit soon thereafter. Thus, while the three trajectories were virtually identical through age 17, they took very different paths of "desistence" at that point.

We will not discuss each of the panels found in Figure 7.8, but we will note one consistent pattern depicted in the panels. The consistent pattern found in the panels of Figure 7.8 is that during the early ages, each of the latent classes (within each panel) is extremely similar to one another and for the most part lie nearly directly on top of one another. However, with the onset of adulthood, the trajectories begin to assume different

[22] It is interesting to note that the LC4 latent class of the 1981-82 sample did not show up in either the 1986-87 sample or the 1991-92 sample. As detailed in Chapter 5, legislation passed in 1982 removed the CYA as a potential sentencing alternative (in place of sentencing to the CDC) for young adult offenders who had been convicted of serious (e.g., index) offenses. Thus, this latent class did disappear from the CYA population, but only because of a change in sentencing patterns. The LC4 latent class most likely just "moved" into the CDC population, rather than disappearing as a type of latent class in the offender population. Nonetheless, the results of the models fit with the changes in the sentencing patterns that removed the LC4 type of offender from the CYA population.

growth patterns. The 1991-92 latent class trajectory in each panel always declines at the fastest rate, while the 1986-87 trajectory falls in the zone in between the 1991-92 and 1981-82 trajectories. These results are entirely consistent with a "period effect" that was serving to modify the shape of the developmental course of the arrest trajectories. The fact that the 1986-87 latent class in each panel is consistently found in the "buffer" zone between the 1981-82 and 1991-92 trajectories provides considerable support to the argument that the change across the samples in each panel is not merely a statistical artifact. A change was occurring over the period of time studied here, and whatever the source of the change (e.g., the declining crime and arrest rates), it was serving to redirect the arrest trajectory on a path of desistence. Relative to the 1981-82 sample, this change occurred faster in the 1986-87 sample, and accelerated even faster in the 1991-92 sample. Again, most important for the concerns addressed here is the fact that the arrest trajectories across the samples were virtually identical all the way through the juvenile portion of the age distribution, and then they assumed different developmental shapes. For example, note the difference in the predicted arrest rates at age 27 between LC4 of the 1991-92 sample and LC2 of the 1981-82 sample (Panel A of Figure 7.8). A finding such as this is simply incompatible with the age-invariance hypothesis and favors the inference that regardless of your prior offending history, behavioral change is possible. Perhaps more importantly, this finding also suggests that broad social conditions can alter trajectories and possibly influence them to either decline faster or slower depending on the nature of the effect that the changing conditions have on the production of crime.

SUMMARY OF RESULTS

Having now completed the presentation of results for this chapter, here we briefly summarize the results presented above. After this review, the chapter concludes with a discussion of these results and how they refute or support the hypotheses detailed in Chapter 3.

The presentation of results in this chapter began with a description of the overall age-crime relationship within each sample. The age crime-curve in all three samples was found to have a rapid increase in adolescence, a peak arrest rate during adolescence, and then a decline in the arrest rate through adulthood. Other than the fact that the peak age at arrest occurred earlier than late adolescence, these results indicate that the aggregate age-crime relationship within each of the samples resembled the overall robust aggregate age-crime curve present in the population at-large (as presented in Figure 1.1).

Next, attention turned to examining whether there were diverse, heterogeneous age-crime trajectories concealed within the overall age-crime curve (computed for all of the sample cases combined). After discussing the method used to arrive at the optimal number of latent classes (e.g., the BIC statistic and local/global testing), we presented the results of the application of the semiparametric mixed Poisson model of Nagin and Land (1993). In all three samples, the 7-class and 8-class models resulted in multiple local maxima, whereas the 6-class model generated a single unique (presumably global) solution. Also, in all three samples, the BIC statistic favored the choice of the 6-class model over the 5-class model. Thus, it was determined that the semiparametric mixed Poisson model with six components or six points of support was the optimal model in all three samples.

At this point, the resulting parameter estimates from each of the 6-class semiparametric models were examined both descriptively and statistically to determine whether the nature of the regression coefficients for the age and age-squared variables were equivalent or *invariant* across the latent classes within each sample. Descriptively, there were discrepancies in the magnitude of the regression coefficients across the latent classes within each sample. For example, the estimated coefficients ranged from, 3.99 to 29.20 in the 1981-82 sample. Next, we statistically tested via the Wald (linear constraint) test statistic whether each set of the regression coefficients were equivalent (within sampling fluctuations) across the latent classes within each sample. The results of the Wald tests indicated a resounding rejection of the null hypothesis that the estimated age parameters (e.g., for the age variable and for the age-squared variable) were equivalent across the latent classes within each sample. These results shed statistical evidentiary doubt on the age invariance hypothesis proposed by Gottfredson and Hirschi.

To more carefully examine whether the statistical differences in the estimated age parameters across the latent classes were substantively important, we next turned to a graphical approach to examine the age-invariance hypothesis. The results were examined on a sample-by sample basis, and our review of the results proceeds accordingly.

Summary of Substantive Results for the 1981-82 Sample

After assigning each individual in the sample to the latent class to which he had the highest (posterior) probability of belonging, a series of descriptive analyses were undertaken. First, the offending patterns of each latent class were summarized. The mean number of total arrest charges was found to vary greatly among the latent classes.

The means ranged from 5.59 average arrest charges all the way up to 44.10 arrest charges, or roughly a difference of 39 arrest charges! The model-based predicted total arrest charges were found to be very similar to the observed average total arrest charges. In each latent class, roughly 50% of their arrest charges were found to be composed of serious offenses. Next the average age at first arrest was described for each sample. The results of the average ages were found to be inconsistent with the age-invariance hypothesis. The age-invariance hypothesis requires the group that has the youngest average age at first arrest to also have the highest mean number of charges. The latent class with the youngest average age at first arrest did not have the highest number of arrest charges.

Next, we presented graphs of the observed and predicted arrest trajectories for each latent class in the 1981-82 sample. The nature of the growth and decline of the arrest trajectories was discussed in terms of whether between-group differences were maintained over time and whether there was an adolescent-limited offender group. An examination of the trajectories indicated that indeed there was an adolescent-limited offender group in the sample. The comparisons of the arrest trajectories in terms of the stability of between-group differences led to a substantive conclusion consistent with the Wald tests—the relationship between age and crime was found to vary across the latent classes. Between-group differences were not maintained over time.

Summary of Substantive Results for the 1986-87 Sample

The same sets of descriptive analyses were then discussed for the 1986-87 sample. The substantive conclusions reached in this sample (and in the 1991-92 sample)

were, in fact, identical to those reached above. Individuals were first assigned to a particular latent class (via the maximum probability rule), and the latent classes were then subjected a series of descriptive analyses. An examination of the offending patterns of each latent class was first summarized. The mean number of total arrest charges was found to be highly variable among the latent classes; the means ranged from 9.46 average arrest charges all the way up to 35.87 arrest charges, or roughly a difference of 27 arrest charges. The model-based predicted total arrest charges were again found to be very similar to the observed average total arrest charges and (again) roughly 50% of the arrest charges in each latent class were found to be composed of serious offenses. Attention then turned to a description of the average age at first arrest. The descriptions of results for the average ages at first arrest were again found to be inconsistent with the age-invariance hypothesis. In this sample, the latent class with the youngest average age at first arrest did not have the highest number of arrest charges, and in fact, ranked third in terms of the average total observed arrest charges (and had nearly 10 less arrest charges than the most frequently arrested latent class).

The observed and predicted arrest trajectories for each latent class in the 1981-82 sample were presented next. The description of the arrest trajectories was focused in terms of whether between-group differences change over time and whether there was an adolescent-limited offender group. An examination of the trajectories indicated the presence of an adolescent-limited offender group in the 1986-87 sample as well. Similar to the results presented for the 1981-82 sample, the comparisons of the arrest trajectories in terms of the stability/unstability of between-group differences led to the substantive conclusion consistent with the Wald tests—the relationship between age and crime was

found not to be invariant across the latent classes. Rather, there were varied substantive differences in the relationship between age and crime found among the various latent classes. In total, the results indicated that between-group differences were not maintained over time.

Summary of Substantive Results for the 1991-92 Sample

The substantive results for the 1991-92 sample were presented next. Building on the consistency for the results presented in the 1981-82 and 1986-87 samples, the substantive conclusions reached for this sample again replicate the results for the two prior samples. Thus, robust results were presented in each of the three samples. Individuals in the 1991-92 sample were again first assigned to the latent class that had the highest probability of having generated the individual's observed longitudinal offense pattern. The offending patterns of each latent class were then described. There was between-class heterogeneity found with respect to the mean number of arrest charges. The mean number of total arrest charges varied from a low of 9.43 arrest charges to a high of 28.57 arrest charges. The groups with the lowest and highest arrest charge totals differed by over 19 arrest charges. For this sample as well, the model-based predicted total arrest charges for each latent class were found to be very consistent with the observed average total arrest charges in the latent class. In each of the latent classes, roughly 50% of the arrest charges were found to be for serious offenses. Turning next to the average age at first arrest, the results were once again found to be inconsistent with the age-invariance hypothesis. The latent class with the youngest average age at first arrest once again failed to accumulate the highest number of arrest charges, and the group

with the oldest average age at first arrest was not the group with the fewest average number of arrest charges.

To further investigate the nature of the between group differences, we presented graphs of the observed and predicted arrest trajectories for each latent class. The trajectories were discussed in terms of whether between-group differences were maintained or changed over time and whether there was evidence of an adolescent-limited offender group. Similar to the two earlier samples, an examination of the observed and predicted arrest trajectories indicated the presence of an adolescent-limited offender group in the 1991-92 sample. Finally, the comparison of the arrest trajectories in terms of the stability of between-group differences similarly led to a substantive conclusion that was consistent with the Wald tests presented at the beginning of the chapter—the relationship between age and crime was again found to *vary* across the latent classes. Differences between the groups were not stable across time, but rather they changed with the waxing of waning of the trajectories of the latent classes.

Summary of the Comparison of Latent Classes Across the Samples

Examination of the trajectories in each of the samples indicated that there were four trajectories common to all three samples, and one trajectory common to only the 1986-87 and 1991-92 samples. Closer examination of the trajectories indicated that the groups were not completely identical across time (which is also incompatible with the age-invariance hypothesis of Gottfredson and Hirschi), and in fact, a pattern was consistently observed in the five panels of Figure 7.8. The consistent pattern seen in Figure 7.8 indicated that after a similar trajectory shape during childhood and

adolescence, the trajectories began to assume different growth patterns with the onset of adulthood. The 1991-92 latent class trajectory in each panel of Figure 7.8 always declined at the fastest rate, while the 1986-87 trajectory fell in between the 1991-92 and 1981-82 arrest trajectories. This result was interpreted as consistent with a possible *period effect* and that broad social conditions can alter arrest trajectories and possibly influence them to decline either faster or slower.

DISCUSSION

The purpose of this chapter was to critically examine the relationship between age and crime among latent classes of serious youthful offenders. This chapter began with a summary description of the different explanations for the shape of the age-crime curve that were presented in Chapter 2 of this study. Recall that Chapter 2 presented a detailed description of the theories of Gottfredson and Hirschi (1990), Sampson and Laub (1993) and Moffitt (1993). The distinction between these theories is critical because each of the theories makes different assertions regarding the stability of individual differences in crime across time, and similarly each theory has a different explanation for the observed aggregate age-crime curve.

Gottfredson and Hirschi (1990) sparked the age crime debate in 1983 with their controversial (and some might argue sociologically blasphemous) age-invariance hypothesis in their article in the *American Journal of Sociology* article entitled, "Age, Crime, and Social Explanation" (Hirschi and Gottfredson 1983). In that article, Gottfredson and Hirschi argued that the relationship between age and crime is "inherent, invariant, and inexplicable" (Tittle and Grasmick 1998)---all people, everywhere, and

within any historical period, tend to commit less crime as they age no matter which source of crime data is used as an indicator of offending. Gottfredson and Hirschi's argue that the shape of the age-crime curve is relatively robust across persons, groups, cultures, and periods. *All* individuals will have their greatest involvement in criminal activity during the late adolescent years of life, and offending declines thereafter. The implication of this argument is that even individuals with vastly different life circumstances, and social, psychological, historical and economic experiences will have similarly shaped age-crime curves across the life course (Greenberg 1985). The key implication of Gottfredson and Hirschi's invariance argument (as specified in their 1990 book) is that the *differences between individuals persist over time*. Group differences in criminal offending histories at any point in time simply reflect group variation in the propensity to commit criminal offenses. Since they posit that criminal propensity once formed is extremely resilient to change, naturally the relationship between age and crime has to be invariant and between-group differences that exist at one point in the age-crime curve must exist at any other point in the age-crime curve.

The age-graded life course theory of Sampson and Laub, on the other hand, specifies the relationship between age and crime as much more variable across individuals and groups. Sampson and Laub (1993, 1997) see the general decline in crime with age as a result of the increasing levels of informal social control that are produced by the salient life events of adulthood, including employment, marriage, and military service. Effective social ties strengthen one's social bond. As adolescents enter adulthood and experience the informal social control that results from their investments in interpersonal relationships such as marriage, parenthood, and work, crime becomes

less likely due to the attachments, involvements and commitments of adult life. Important for our concerns here, is that Sampson and Laub invoke a state dependence argument that allows for criminal propensity to be variable over time. Thus differences in criminal propensity are not necessarily stable across time. In fact, given their focus on the sources of informal social control that arise during adulthood, they stress that adulthood is the precise time when preexisting individual differences become less relevant. Rather, it is more important that individuals experience the strengthening of the social bond that often accompanies movement into the various adult roles and responsibilities (Cernkovich and Giordano 2001: 372). Sampson and Laub have been vocal critics of both the age invariance argument of Gottfredson and Hirschi and the presumption of stable individual differences in the propensity to commit criminal acts across the entire life course.

Finally, Moffitt's theory hypothesizes the existence of two (and only two) distinct groups of offenders. According to her, the aggregate age-crime relationship takes on its observed shape because the two discrete offender groups are mixed in the population at-large. The upward surge of the age-crime curve is the result of increasing participation rates of the adolescent-limited offender group, whereas the downward surge results from the termination of offending by this group. Since the adolescent-limited offender group outnumbers the life-course-persistent group (who are hypothesized to commit criminal and antisocial acts at a relatively constant rate across the life course), the offending patterns of the adolescent-limited group are argued to determine the shape of the curve. The life-course persistent offenders account for the offenders in the childhood and adulthood tails of the curve.

In Chapter 3, we reviewed the extant literature on studies of the age-crime curve (within homogenous latent classes of offenders) and concluded there that several current limitations with the previous studies on this topic that necessitate further research on this topic. Two key limitations were discussed in Chapter 3. First, only one study to date (by D'Unger et al. 1998) has examined results from more than one dataset that is generalizable to the same population over time. This limitation makes it difficult to replicate not only the existence of a crime trajectory group over time, but also to establish whether there are any changes in the precise number or nature of the offending trajectories over time. Second, there have been only two studies of the age-crime curve within samples of high-risk offenders (Laub et al. 1998; Piquero et al. 2001), and these studies have one or more of limitations necessitating continued research (e.g., use of nonrandom samples, data limited to white juveniles sampled in the 1930s, limited segment of age distribution studied). Three key questions were noted as understudied in the extant literature with respect to the serious youthful offender population: (1) how many "latent classes" of offenders are necessary and sufficient to capture the variation of offending trajectories in the serious youthful offender population; (2) how do differences in offending trajectories during the juvenile years relate to the nature of offending during the adult years (i.e., are the between-group differences maintained over time); (3) Is there an adolescent-peaked group within this population?

With these questions in mind, this study set out to investigate four hypotheses related to the age-crime curve using three large, random samples of serious youthful offenders. The first two hypotheses noted in Chapter 3 were:

$\mathbf{H_1}$: There are multiple groups or latent classes of offenders with distinct offending trajectories even on the high-end of the criminal propensity continuum where the serious youthful offenders are located.

$\mathbf{H_2}$: There are more than two groups of offenders with distinctly different trajectories even on the high-end of the criminal propensity continuum.

The semiparametric mixed Poisson model of Nagin and Land (1993) was used to empirically tease out the latent classes in the three samples. After determining whether a unique (presumed global) solution could be obtained, the BIC statistic was used as a statistical guide for determining the optimal number of latent classes present in the data. In all three samples, the BIC statistic favored the finite mixture model with six components in the mixing distribution. Thus, the results for all three samples indicated significant support for the first hypothesis, $\mathbf{H_1}$, that there are multiple, distinct offender groups on the high-end of the criminal propensity continuum. These results support the previous descriptive (rather than empirical) contentions of Cohen and Vila (1996) and D'Unger et al. (1998) that the far end of the continuum has far greater heterogeneity than previously thought.

These results also provide evidence refuting the claims of Moffitt (1993) that there are only two discrete offender groups (i.e., the evidence refutes hypothesis $\mathbf{H_2}$). The BIC statistics in all three samples favored a 6-class model. Furthermore, the positive increase in the magnitude of the BIC statistics between the models that only allow for two latent classes (corresponding to the hypothesized number of discrete groups in the

Moffitt theory) and those that allow for six latent classes were indeed quite large (3932.2 in the 1981-82; 1920.5 in the 1986-87 sample; and 1343.9 in the 1991-92 sample).

Further support for the notion that there are more than two groups in the offender population was provided after assigning each individual to the latent class to which they had the greatest probability of belonging. Examination of both the observed average total arrest charges and the observed and predicted arrest trajectories in each latent class indicated that there was simply too much heterogeneity in the population (both in terms of the mean rates of offending and the developmental shapes of the arrest trajectories) to be adequately and sufficiently accounted for with only two latent classes. Examination of the latent class parameter estimates indicated further evidence refuting the dual taxonomy theory of Moffitt (1993). No latent class in any of the samples was found to offend at a relatively constant rate across the age distribution in the hypothesized "spirit" of the life-course-persistent offender group—the age and age-squared parameter estimates for all of the latent classes were found to be significantly different from zero (and indicative of a quadratic relationship). This contradicts Moffitt's (1993: 695) explicit contention that age is not a predictor of the offending trajectory in the "life-course-persistent" group. Further, the graphical results presented in this chapter also failed to lend support to the existence of an offender group that offends across the age span at a relatively constant and persistent rate independent of age. This was found to be true even in this select group of serious youthful offenders where presumably, if there were such a life-course persist-group, it should have been identified.

The third hypothesis examined in this chapter was:

H_3: There is an adolescence-peaked group even in samples of serious youthful offenders.

The examination of the predicted and observed arrest trajectories provided overwhelming support for the presence of an adolescent-limited offender group in the serious youthful offender population. In all three samples, a latent class offender group was identified that was clearly arrested in an "adolescent-limited" pattern. Importantly, for the final 16 age-years in the 1981-82 sample, the final 11 age-years in the 1986-87 sample, and the final 9 age-years in 1991-92 sample, not a single individual assigned to the adolescent-limited offender groups was arrested for even a single charge. This is an impressive finding given that 10%, 13%, and 11% of the 1981-82, 1986-87, and 1991-92 samples, respectively, were assigned to this offender group, and is extremely notable given the proclivity that the members of these samples have shown for getting arrested. This group also poses the most trouble for the next hypothesis that was studied herein:

H_4: The age-crime curve is invariant among the latent classes of serious youthful offenders. Between-group differences will not vary across time.

The results presented in this chapter *for all three samples* send a vigorous signal indicating a lack of support for the H_4 hypothesis. The age invariance hypothesis was first statistically tested using the (linear constraint) Wald statistic that tests the restriction of constraining each age parameter to be equivalent across the latent classes. The

statistical evidence strongly rejected the null hypothesis of no difference in the estimated age parameters across the latent classes within all three samples.

Next, the age-invariance hypothesis was tested by examining the stability/instability of between-group differences in terms of the observed and predicted arrest trajectories of the latent classes. The results in all three samples provided resounding evidence of a breakdown in the maintenance of between-group differences across time. In all three of the samples, the maintenance of between-group differences in arrests was relatively strong only through mid-adolescence. Soon thereafter, however, there was a considerable amount of change in the between-group differences throughout the remainder of the adult years studied here. Indeed, the mere presence of the adolescent-limited offender group poses an absolutely insurmountable hurdle for the age-invariance hypothesis, especially since this latent class was shown to have the lowest incarceration rates in adult prison and because mortality data was used to exclude dead individuals from being considered in the at-risk population at ages subsequent to their death. Therefore, their rapid decline cannot be argued to be simply a consequence of the differential effects of incapacitation and/or mortality. Thus, the group that at one point consistently had one of the highest arrest rates (at around age 15-16), just several years later had the lowest arrest rate. Thus, arrest records indicate that this group had indeed, for all intents and purposes, terminated their offending (in terms of arrest activity at least). This finding is completely incompatible with the hypothesis that the relationship between age and crime is invariant. It is important to remember that this finding was documented across three separate samples, which poses a considerable problem for any rival hypothesis that this pattern was an anomaly or a statistical fluke.

Further, comparisons across the samples in terms of the similarity of the arrest trajectories of the latent classes indicated that there were four robust offender trajectories discovered among all three samples, and another trajectory was discovered to be present in both the 1986-87 and 1991-92 samples. However, even in the presence of a similarity of offender groups across the samples, the trajectories were not identical in all three samples. In fact a consistent pattern was uncovered that appeared to indicate a possible period effect (or perhaps some other consistent source of change) that was causing the arrest trajectory to decrease faster in the 1991-92 sample than it did in the 1981-82 sample. Regardless of the actual cause of the changing shape of the trajectories (within the groups that share a similar trajectory), we found that there were offender groups that (for a significant portion of the age distribution) had nearly identical arrest trajectories. These same groups later had arrest trajectories that were no longer identical. This finding lends further substantive support to the notion that between-group differences (or in this case "between-group similarities") are variable over time. Indeed, the findings in this chapter resonate extremely well with the earliest study to examine the relationship between and crime within discrete latent classes of offenders: "explanations of the age-crime curve are not easily reduced to summary statements about average population tendencies" (Nagin and Land 1993: 358). According to the results presented in this chapter, this appears to be the case even within the serious youthful offender population.

Before concluding this chapter, several final comments are in order. First, it is clear from the results presented here that longitudinal data are absolutely necessary for examining the causes of crime. Lacking longitudinal data, one would lose sight of the fact that even within these samples of persistent offenders, there exists a group of

offenders who only appear to have high rates of criminal activity during adolescence. In fact, if you took a cross-section of these offenders at around age 15, you would find that the adolescent-limited offenders appear to be among the highest rate offenders in the sample. However, if you took the cross-section during their early twenties, these same offenders would be the lowest rate offenders (at a rate near or equal to zero) in the three samples. Of course, the connection that these are the same individuals at two distinct points in time could only be deduced with longitudinal data. Gottfredson and Hirschi (1987) were correct in their contention that there are high costs associated with the collection of longitudinal data (which the authors of this study know all too well), but we would argue it is clear from the results presented in this chapter that an adequate and full understanding of criminal behavior of these individuals could never be accomplished through the use of cross-sectional data only. For example, many criminologists will probably be surprised by the existence of an adolescent-limited offender group within the CYA offender population, given that the CYA wards are renowned for their excessive failure rates upon exit from the CYA institutions and their persistence in offending through adulthood (see e.g., Haapanen 1990; Piquero et al. 2001). Without longitudinal data, one would completely lose sight of the fact that even in the population of high-rate serious offenders, there is an adolescent- limited offender group. Of course, as Tittle (1988: 76) noted, "whether longitudinal data are preferred over cross-sectional is something like asking whether hammers or saws are more useful to carpenters." Cross-sectional data have their strengths and weaknesses, as do longitudinal data, but favoring one at the complete expense of the other would be a serious error equal only to deciding

that only quantitative or qualitative analyses should be undertaken to best understand criminal behavior.

Second, the findings in this study clearly indicate a significant amount of heterogeneity in the propensity to offend within this population, a fact that is important for both theoretical and public policy reasons. The results here indicate that it is dangerous to think of this population as being "relatively homogenous" (Ge et al. 2001: 750). As a whole, serious youthful offenders are an elusive class of offenders because they are rare in the population of offenders (Cernkovich et al. 1985). Researchers, however, should keep in mind that even within this segment of the population there is a considerable amount of heterogeneity. The offenders from the serious youthful offender population undoubtedly will always stand out when (assuming they are actually sampled by chance) they are found in general population samples. As the results here clearly indicate, just because these offenders "stand out" compared to non-offenders or low-rate offenders in the general population, does not mean they then "stand together" when you actually examine the longitudinal offending patterns of a large sample of such offenders. There simply is much more heterogeneity in this population than has previously been acknowledged.

Finally, the results obtained in this study suggest the limited utility of the dual offender typology employed by Moffitt. This is especially troubling because of the tendency in much contemporary criminological research to investigate the tenets of this theoretical perspective after dividing the sample into two groups (which are then labeled LCP and AL) on the basis of age of onset alone (see e.g., Dean et al. 1996; Piquero et al. 1999, Scholte 1999; Aguilar et al. 2000; Klevens et al. 2000; Mazerolle et al. 2000,

Cernkovich and Giordano 2001; Ge et al. 2001; Piquero and Brezina 2001). The results presented in this study indicate that it was not the age of onset that differentiated the adolescent-limited offender group from the other offender groups, but rather it was the unique developmental nature of their arrest trajectory. Given that there appears to be more offender groups in the population than simply two (and that age of onset appears to be a questionable method of separating out the two groups), analyses and interpretations based on this dual taxonomy distinction might appear to be a helpful heuristic device, but in practice they may be: 1) of questionable theoretical import and 2) potentially misleading. If populations/samples/datasets cannot be neatly and discretely divided into two groups by arbitrarily dividing them on the basis of age of onset (and the results obtained herein indicate that they can not), then such a process is likely to do nothing other than reify the dual offender categories as if they actually exist in the offender population.

CHAPTER 8

ON THE RELATIONSHIP OF PAST TO SUBSEQUENT CRIMINAL ACTIVITY

INTRODUCTION

Here, we first briefly review the theoretical importance of studying the relationship between past and subsequent criminal activity before proceeding with a presentation and discussion of the data analyses conducted on this topic in this chapter. Just as the last chapter critically examined one of the most robust findings in criminology—the supposed invariance in the age-crime relationship, here the relationship between past and subsequent criminal activity (another robust finding in criminology) will be explored. As Brame et al. (1999: 600) note:

> The strong positive association between past and subsequent criminal offending is one of the most agreed, yet least well understood facts about law breaking behavior. Individuals who have offended in the past are most likely to offend in the future. There is little doubt or ambiguity about the validity of this claim. *Still, it is not clear why this association exists* (emphasis added).

In other words, the fact that there is a positive association between criminal offending at two (or more) points in time is really not in question; what is at issue, however, is the etiological significance of this association. As presented in explicit detail in Chapter 2, there are three broad etiological expositions that assert unique alternative explanations for this recurrently documented positive association: (1) population heterogeneity; (2) state dependence; and (3) dual taxonomy "mixed" theories of crime.

Gottfredson and Hirschi's (1990) general theory of crime asserts that the association between past subsequent criminal activities is spuriously due to population

heterogeneity in the propensity to offend. After properly controlling for individual differences in the propensity to commit criminal acts, the association between these variables should be reduced to the immediate region of the null hypothesis--zero. In the absence of proper controls for individual differences, the association between past and subsequent criminal activities exists because those with high criminal propensity consistently offend in adjacent measurement periods, which naturally (and spuriously) induces an association between criminal offending at any two points in time.

Sampson and Laub's (1993) theory of age-graded informal social control was described as an example of the state dependence explanation of the association of past and subsequent criminal activity.[1] Sampson and Laub's theory posits that there will be a significant positive association between past and subsequent criminal activity, even after controlling for persistent differences in the propensity to offend, because criminal activity serves to "knife off" opportunities for prosocial activities and makes continuing in a lifestyle of crime more likely. Stated differently, committing crimes (and being arrested) has deleterious consequences on the "local life circumstances" of an offender, thereby making future crime more likely (Horney et al. 1995).

Finally, Moffitt's (1993) dual taxonomy theory was used as an example of a theory that incorporates both the population heterogeneity and state dependence arguments into its theoretical exposition. According to Moffitt, the theoretical framework governing the criminal behavior of the "life-course-persistent" offender is a static population heterogeneity process that has run its course by the end of childhood,

[1] As reviewed in Chapter 2, the theory of Sampson and Laub (1993) is not a "pure" state dependence theory because they do recognize the empirical and theoretical importance of individual differences in criminal propensity. However, the major theoretical thrust of their argument is a state dependence explanation, which is why we have characterized it as such in this study

whereas the state dependence explanation governs the offending patterns of a different " adolescent-limited" group of offenders. After empirically separating the two distinct groups, researchers should find: (1) no relationship (or a severely reduced effect at most) between past and subsequent offending *within* the life-course-persistent group (periods of criminal activity are followed by further criminal acts merely because of their time invariant high-levels of criminal propensity); and (2) a strong, positive association between the offending patterns in the adolescent-limited group (there is a strong *causal*, state dependence effect resulting from the positive reinforcement contingencies of achieving mature status with the criminal acts). Of course the results presented in the last chapter cast empirical doubt on the claims that (1) there are only two offender groups in the population, and (2) that there is a group of offenders who commit criminal acts *persistently across the entire life course*. But still, an adolescent-limited group of offenders was found in all three of our samples, and thus the empirical question of the importance of the relationship of past to subsequent criminal offending *within* the adolescent-limited group is still a significant issue deserving of empirical investigation. The results we present in this chapter to our knowledge, represent the first empirical attempt to examine the prominence of state dependence processes within a group of offenders shown to have acted in an adolescent-limited pattern.

Our review of the extant literature on this topic in chapter 3 indicated that there are several current limitations in the previous literature highlighting the need for additional analyses such as that undertaken here. There were two key limitations pointed out in Chapter 3. First, there are questions regarding the validity of the observed state dependence effects identified in prior studies due to the possible consequences of

violations in the assumptions of the statistical models employed in these prior studies. Second, there are also questions regarding whether population heterogeneity processes "rule the day" in high-risk samples such as those employed in our study.

In view of these two key limitations in previous research, we now present a series of analyses that aim to address whether there are any state dependence effects within the three samples of what can only be described as "very high-risk" offenders documented by the evidence of offending patterns presented in Chapter 6. The fact that we have "very high-risk" samples allows us to examine the presence of state dependence effects within this segment of the offender population. In addition, the application of several different analytical methods allows us to assess whether the results are robust to the specific analytical approach employed to control for persistent individual differences.

This chapter has three main sections, with one main section of results dedicated to each of the three release samples. In this chapter we employ the use of the multimethod approach of Bushway et al. (1999), which is essentially the "compare and contrast strategy" recommended by Heckman and Singer (1984). More specifically, we will test the robustness of any observed effect of past and subsequent criminal behavior by employing several different methods of analysis.

As described in the final section of Chapter 5, we undertake five stages of analysis for each release sample. In the first stage, we employ the use of the Poisson finite mixture models to estimate the magnitude of the state dependence parameter while nonparametrically controlling for the unobserved heterogeneity. Stage two contains a presentation of the results from the parametric random effects model, where the unobserved heterogeneity is assumed to follow a specific parametric distribution (the beta

distribution) and where the age parameters are assumed to apply equally to all individuals. In stage three, we incorporate the latent class indicators from the results presented in Chapter 7 in order to address two questions. First, does the set of latent class indicators remove the presence of unobserved heterogeneity? This question will be answered by comparing the results of the parametric random effects model that includes the latent class indicators to the results of the NB1 model with the latent class indicators. Second, does allowing the estimates of the age parameters to vary over the latent classes improve the estimation of the state dependence parameter? Stage four examines the effect of the state dependence parameter estimate calculated within each of the latent classes by themselves. Our analyses conclude with stage five, in which we examine whether (1) the effects observed in this study would have changed if only the post-release arrest data were available and (2) if there are any covariates significantly related to the post-release arrest rates, net of the effects of unobserved heterogeneity. The results of our analyses will be presented separately for the 1981-82 sample, the 1986-87 sample, and finally for the 1991-92 sample, respectively. It bears repeating that differentiating state dependence processes from population heterogeneity processes with respect to the relationship between past and subsequent criminal activity is a methodologically complicated task, a point that should not be underemphasized (Nagin and Paternoster 2000). We have included technical comments and footnotes when necessary and relevant in the remaining portions of this chapter.

RESULTS FOR THE 1981-82 SAMPLE

For the 1981-82 sample, we describe each analytical stage in detail as well as the rationale behind its use. The analyses that will follow for the 1986-87 and 1991-92 samples are mere replications of the same analytical processes employed here for the 1981-82 sample, and thus will not described in the same detail.

Stage One—The Semiparametric Mixed Poisson Model

We begin our analyses by employing the semiparametric mixed Poisson model of Nagin and Land (1993; Land and Nagin 1996; Land et al. 1996; Nagin 1999). Again, this model controls for persistent individual differences through the nonparametric "points of support" approach. The only assumption this model makes with respect to the distribution of unobserved heterogeneity is that it can be approximated using a discrete, multinomial distribution. The random effects, in other words, are assumed to have been drawn from $j = 1, 2, 3, \ldots, J$ discrete groups. This model is an excellent model to begin the presentation of results for two reasons. First, a model with one point of support corresponds, in fact, to a standard Poisson model that assumes no stochastic variation and no unobserved heterogeneity, and thus provides an excellent baseline model for subsequent comparisons. Second, by specifying models with increasing numbers of points of support, we can numerically observe the effects that improved control for unobserved heterogeneity has on the various magnitudes of the state dependence effects. For example, we can compare the estimated state dependence effect in a two points of support model (or 2-class model) with the estimate from a three points of support (or 3-class) model.

The specification used in the semiparametric models in this chapter is similar to the specification of these models employed in Chapter 7, with the exception that the model specifications in this chapter also include a binary indicator of arrest in the prior period, (arr_{t-1}). The formal specification employed in this chapter is:

$$\ln\left(\lambda_{it}^{j}\right) = \left(\beta_0 + \bar{\varepsilon}_j\right) + \left((age_{it}/10) * \beta_{age}^{j}\right) + \left(\left(\left(age_{it}^2\right)/100\right) * \beta_{age^2}^{j}\right) + \left(arr_{t-1} * \beta_{arr_{t-1}}\right). \qquad (1)$$

This specification allows for both latent class-specific (or group-specific) intercepts and age coefficients, but the effect of the regression coefficient that estimates the state dependence relationship between the binary variable (indicating arrest in the prior period) and the mean offending rate in the current period, denoted as $\hat{\beta}_{arr_{t-1}}$ in equation (1) above, was constrained to be equal across the latent classes. This assumption is relaxed in stage four below.

The same process for fitting and testing the semiparametric mixed Poisson model with varying numbers of "points of support" described in Chapter 7 was also employed here. This model-fitting process uses both the BIC statistic and the testing for global/local solutions. The testing for local/global solutions was identical to that used in Chapter 7. Models with more than one solution (i.e., local solutions) were identified through this extensive model testing procedure and removed from consideration.

Similar to the results presented in Chapter 7, the 7-class and 8-class models in the 1981-82 sample were again prone to local solutions that varied from one solution to another. The 6-class model, however, generated the same unique solution all ten times it

was estimated. Further, and as shown in Table 8.1, the 6-class model had the largest BIC value, and thus this model was determined to be the model with the optimal number of latent classes. The 6-class model had a BIC value of -72391.62, whereas the BIC value for the 5-class model was only -72670.64. However, since interest in this chapter also lies in changes in the magnitude of the state dependence coefficient, we also present the solutions for all the models up through the 6-class model in Table 8.1. Thus, the optimal number of latent classes in this chapter was used to end the presentation of model results (i.e., results from the 7-class solutions are not presented), rather than to present only the solutions from that model.

Table 8.1 presents the results of the 1- through 6-class semiparametric mixed Poisson models. Interest here focuses exclusively on the regression coefficient, $\hat{\beta}_{arr_{t-1}}$, found in the row identified as, "Arr_{t-1}." This estimate represents the state dependence relationship between the binary variable indicating arrest in the prior period (arr_{t-1}) and the mean offending rate in the current period (age). The first column in Table 8.1 contains the parameter estimates from the 1-class model. The estimate of the state dependence parameter was 0.857, which was highly significant with a t-statistic of 86.28 (which supports the fifth hypothesis of this study).[2] Again, this model corresponds to a standard Poisson regression model with no stochastic variation and under the assumption

[2] For comparative purposes, a t-statistic value of 1.96 is significant at the .05 level, a value of 2.58 is significant at the .01 level, and a value of 3.30 is significant at the .001. These values (.05, .01, and .001) are the conventional "levels of significance" used in most empirical research.

Table 8.1. Investigation of State Dependence Effects with Semiparametric Random Effects Poisson Model: 1981-82 Sample (N = 1989; Panel Observations = 60453)

	Points of Support					
	1	2	3	4	5	6
Random Effects						
Class1	-2.978	-3.015	-3.867	-4.994	-4.986	-4.663
	(74.58)	(60.35)	(46.80)	(39.59)	(38.09)	(35.58)
Class 2	---	-5.560	-2.911	-3.523	-2.756	-2.7[illegible]3
		(38.81)	(36.26)	(22.81)	(32.17)	(30.18)
Class 3	---	---	-15.229	-3.067	-5.603	-5.374
			(24.63)	(26.99)	(26.68)	(28.23)
Class 4	---	---	---	-14.855	-19.289	-21.242
				(18.71)	(23.35)	(21.75)
Class 5	---	---	---	---	-3.771	-8.849
					(29.59)	(14.55)
Class 6	---	---	---	---	---	-3.689
						(28.17)
Age Effects						
Class 1 Age	2.511	2.664	3.715	3.962	3.903	3.716
	(64.71)	(57.98)	(43.78)	(36.41)	(34.77)	(33.59)
Class 1 Age-Squared	-0.585	-0.577	-0.923	-0.842	-0.814	-0.766
	(67.92)	(58.48)	(42.61)	(35.33)	(34.33)	(32.99)
Class 2 Age	---	5.943	2.625	4.059	3.067	3.103
		(34.88)	(39.32)	(24.41)	(36.27)	(34.65)
Class 2 Age-Squared	---	-1.657	-0.544	-1.083	-0.752	-0.762
		(33.77)	(39.91)	(23.05)	(37.45)	(36.07)
Class 3 Age	---	---	19.056	2.764	6.430	6.221
			(23.65)	(32.97)	(27.90)	(29.06)
Class 3 Age-Squared	---	---	-5.935	-0.555	-1.780	-1.722
			(22.87)	(35.46)	(28.37)	(29.30)
Class 4 Age	---	---	---	18.535	24.515	28.031
				(17.57)	(22.88)	(21.67)
Class 4 Age-Squared	---	---	---	-5.743	-7.727	-9.059
				(16.67)	(22.43)	(21.29)
Class 5 Age	---	---	---	---	3.152	7.784
					(30.03)	(12.17)
Class 5 Age-Squared	---	---	---	---	-0.586	-1.896
					(27.91)	(11.29)
Class 6 Age	---	---	---	---	---	3.127
						(28.75)
Class 6 Age-Squared	---	---	---	---	---	-0.579
						(26.48)
State Depen. Effects						
Arr_{t-1}	0.857	0.618	0.545	0.470	0.439	0.404
	(86.28)	(61.07)	(53.43)	(45.50)	(42.38)	(38.82)
Log-Likelihood	-78225.72	-75155.540	-73945.49	-73047.13	-72594.68	-72300.470
BIC	-78240.91	-75185.92	-73991.06	-73107.9	-72670.64	-72391.62

Note: Absolute values of t-statistics are in parentheses.

of complete homogeneity with respect to criminal propensity (which as was already shown in Chapter 7 to be an erroneous assumption).[3]

The next column in Table 8.1 presents the results from a two-class model that assumes there are two latent classes in the population. The resulting parameter estimate of the state dependence relationship in this model was 0.618 (t-statistic = 61.07), which, in comparison with the estimate from the one-class model, represents a 27.9% decrease in the magnitude of the parameter estimate.[4] The next column corresponds to a 3-class model (which allows an 3 additional parameters—an intercept and two additional age parameters), and here again there was a decline in the magnitude of the state dependence parameter estimate from 0.618 (2-class) to 0.545 in this model. The estimate was still highly significant in this model with an estimated t-statistic of 53.43. Allowing for four points of supports further reduced the parameter estimate to 0.470 (t-statistic = 45.50), and allowing for five points of support reduced it even further to 0.439 (t-statistic = 42.38). Lastly, the final state dependence estimate in the 1981-82 sample was 0.404 when we allowed for six points of support (which can be found in the last column in Table 8.1).

---

[3] For comparative purposes, we also estimated a standard negative binomial regression model (NB1) that at least removes the assumption of a lack of stochastic variation. Allowing for the stochastic variation should provide a significant increase in the standard errors, a significant reduction in the corresponding t-statistics, but very little change in the magnitude of the state dependence parameter because individual differences are still left uncontrolled in this model. As expected, the parameter estimate in the NB1 model was .871, but the t-statistic decreased to 58.80 due to the substantial increase in the standard errors in this model.

[4] Ideally, one would like to conduct a formal statistical test concerning whether the difference is equal to 0. However, since the statistical theory underlying such a test presumes that the samples on which the parameter estimates are calculated are independent, there is no formal test available (Paternoster et al. 1997). The problem is that the covariance between the two estimated coefficients is unknown and it cannot be assumed to be zero (since the samples on which they are both estimated are the exact same samples)

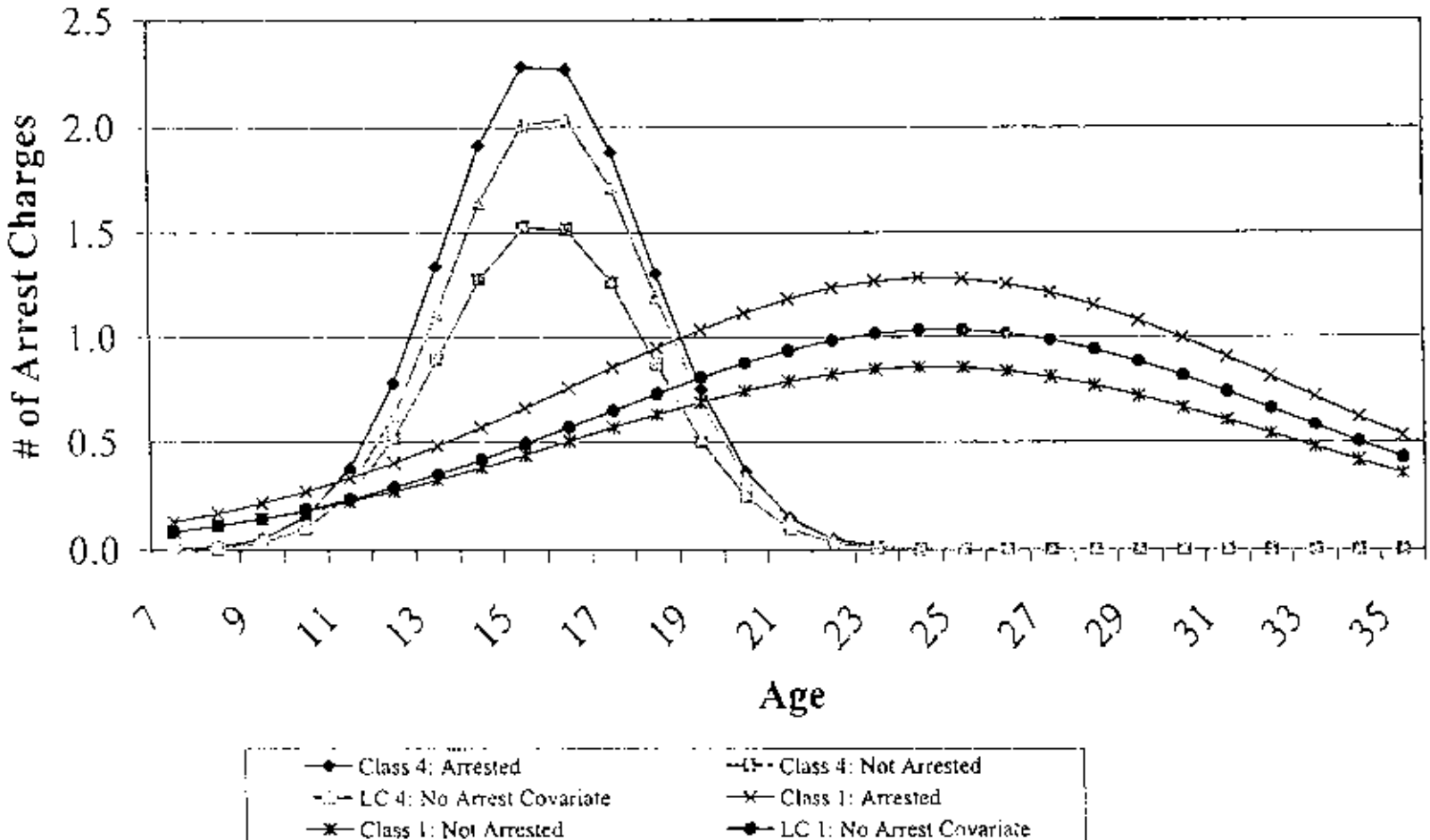
Three items at this point are important to note. First the state dependence parameter is positive, and well above zero even in the 6-class model. The positive parameter estimate in the final model indicates that even after controlling persistent individual differences (through the points of support approach), individuals who were arrested at a prior age had a higher mean number of arrest charges at the next age in comparison with the individuals who were not arrested at the immediately prior age. Second, the state dependence parameter estimate is not only positive, but it is still highly significant with an associated t-value of 38.82 in the 6-class model. In other words, having been arrested at a prior age *significantly* increased the rate of criminal activity at the next age, a finding that is entirely consistent with the state dependence position of Sampson and Laub (1993) and incompatible with the pure population heterogeneity explanation espoused by Gottfredson and Hirschi (1990). This result supports the seventh hypothesis examined in this study, and refutes our sixth hypothesis. Third, controlling for persistent individual differences was critical in terms of calculating the precise numerical estimate of the state dependence parameter. In the standard Poisson model (1 point of support model), the parameter estimate was 0.86, whereas controlling for the significant individual differences reduced this estimate to 0.40, which amounts to a 53% reduction in the absolute magnitude of the effect. Clearly, the results for the 1981-82 sample derived from this model lend support to the "mixed" model approach used to explain the relationship between past and subsequent criminal activity. Both population heterogeneity and state dependence explanations appear to make significant contributions toward explaining the relationship between criminal offending at two age periods.

As noted in Chapter 7, the semiparametric mixed Poisson model contains a number of nonlinear terms that make it difficult to substantively interpret the parameter estimates from this model. To aid in understanding the substantive importance of the state dependence parameter estimate in the 6-class model found in Table 8.1, we generated predicted arrest trajectories for two of the latent classes in this model. Figure 8.1 presents the predicted offending trajectories for the first and fourth latent classes in the 6-class model in Table 8.1. This figure contains six predicted trajectories, three for each class. Two of the trajectories represent the predicted trajectories for each of the two latent classes that were presented in Chapter 7 in which there was no covariate in the model indicating arrest at the prior age (these two trajectories are labeled "No Arrest Covariate" in Figure 8.1).[5] The other two trajectories for each latent class that are presented in Figure 8.1 represent the predicted mean arrest charge for cases who (1) were arrested at the prior age (labeled "Arrested" in Figure 8.1) and (2) were not arrested at the prior age (labeled "Not Arrested"). The two trajectories "borrowed" from the results presented in Chapter 7 lie in between the other two trajectories for each latent class because the effect of having been arrested at a prior age was "averaged" out in the model presented in Chapter 7 (which did not have a covariate controlling for this effect).

Nonetheless, the important point of Figure 8.1 is that not only is there a statistically significant difference between the mean arrest rates of those who were and

[5] These two predicted trajectories are the exact trajectories presented in Chapter 7. Note that in Chapter 7 the adolescent-peaked trajectory in the 1981-82 sample was labeled "LC5" because it was the latent class that ranked 5th in prevalence. This adolescent peaked trajectory in this chapter was the fourth latent class in Table 8.1 because it ranked 4th in prevalence (i.e., with respect to how many people are estimated to belong to the latent class). In Chapter 7, "LC5" had 215 cases (10.81%) assigned to it, whereas in the 6-class model of Table 8.1, the adolescent peaked group had 222 cases (11%) assigned to it. The 7 additional cases were enough to make it rank fourth in prevalence.

Figure 8.1 Predicted Mean Arrest Charges, by Arrest Status at Prior Age: 1981-82 Sample



were not arrested at a prior age (i.e., the difference is not zero), but more importantly there is a *substantively important difference* in their mean offending rates as well. For example, for the adolescent-limited trajectory in Figure 8.1, there is a difference (between those who were and were not arrested at the prior age) of about 0.75 arrest charges (on average) between the ages of 15 and 18. Again, recall that the descriptive presentation of the arrest histories of these individuals in Chapter 6 indicated that over half of all of their arrest charges were for serious criminal offenses, and thus this difference is substantively meaningful. The majority of the criminal offenses for which these individuals were arrested were not trivial matters, and thus differences between mean offending rates have extremely important implications in terms of the societal costs.

Before moving on to stage two of the analysis, it is interesting to note that the decrease in the magnitude of the state dependence parameter estimate that occurs with each additional point of support became smaller as we added additional points of support to the model. For example, the addition of a second class resulted in a decrease of 0.239 in the state dependence parameter estimate (i.e., 0.857 - 0.618 = 0.239), whereas going from the 3-class model to a 4-class model resulted in a decrease of 0.075, and going from a 5- to 6-class model resulted in a change of only 0.035 in the parameter estimate. These results are entirely consistent with how the "point of support" methods approximate the mixing distribution. Eventually the extraction of an additional point of support results in two or more of the points of support becoming similar, which is what eventually results in the model failing to reach convergence (or leading to local solutions with weak identification). Mathematically, the new additional point of support becomes too similar to one of the other points of support, and this causes the model to "blow-up" because the

Hessian matrix becomes singular (i.e., two columns of the matrix are linearly dependent) and cannot be inverted.[6] Substantively, however, the fact that the change in the state dependence parameter estimate became smaller as the higher-order points of support were added to the model was a clear indication that the heterogeneity in the mixing distribution was accurately approximated with the finite number of points of support.

Stage Two—The Parametric Random Effects Negative Binomial Model

Attention is now turned toward the typical approach employed in previous research to control for individual differences—the parametric random effects model. In this stage, we employ the use of the parametric random effects negative binomial model and estimate the following model:

$$\ln(\lambda_{it}) = \beta_0 + \left((age_{it}/10) * \beta_{age}\right) + \left(\left((age_{it}^2)/100\right) * \beta_{age^2}\right) + \left(arr_{t-1} * \beta_{arr_{t-1}}\right) + \varepsilon_{it}. \qquad (2)$$

[6] Wedel and Kamakura (1998) found that this is also one of the principal causes of local solutions. The local solution models differ from non-convergence models in that the local solution models are still able to obtain estimates (i.e., the Hessian matrix can still be inverted and standard errors can still be calculated), but they indicate a "weak identification" of the model since the "points of support" are not clearly separated or well-defined. Indeed, our investigation of the local solutions in the release samples analyzed in this chapter indicated that this is what was precisely occurring in these cases. The local solutions were "weakly identified" meaning that two or more of the points of support were not very different from one another. In these local solutions, the model was still able to invert the Hessian matrix because the columns of this matrix were not so linearly dependent as to cause the determinant of the matrix to become zero (which results in the matrix being uninvertible). It is important to note that in these "local solution" models, the estimate of the state dependence parameter hovered in the general vicinity of the estimated parameter estimate from the 6-class model. This lends further support to the conclusion that the 6-class models described in this chapter are sufficient to account for the distribution of unobserved heterogeneity

Again, we are interested in the regression coefficient, $\hat{\beta}_{arr_{t-1}}$, that estimates the state dependence relationship between the binary variable indicating arrest in the prior period (arr_{t-1}) and the mean offending rate. Unlike the semiparametric model, this model assumes the distribution of unobserved heterogeneity follows a very specific, mathematically tractable parametric distribution, which in the current model is the beta distribution. To the degree that the unobserved heterogeneity is distributed in the population according to the assumed distribution, the random effects model is more efficient than the semiparametric model. The random effects models have arguably been used more frequently to control for unobserved heterogeneity because these models have been readily available in canned software packages commonly used by social scientists (e.g., SAS, Stata, LIMDEP). The complexity of the finite mixture models and a lack of available software have certainly limited the use of these models to control for unobserved heterogeneity. Here we will compare the estimates of the state dependence effect from this model using the parametric specification of the unobserved heterogeneity to those obtained in the first stage using the nonparametric specification of the random effects.

The results from the random effects negative binomial model are found in the first column of Table 8.2 (under the Model 1 column heading). This model involves the estimation of four regression coefficients—an intercept, age, age-squared, and the state dependence parameter. To test for the presence of significant unobserved heterogeneity, we performed a "boundary-value likelihood ratio" that compares the random effects negative binomial model against the NB1 negative binomial model (Gutierrez et al.

Table 8.2. Investigation of State Dependence Effects with Parametric Random Effects & NB1 Negative Binomial Models: 1981-82 Sample

	Model			
	1	2	3	4
Overall Intercept	-3.834	-4.590	-3.993	-17.869
	(65.13)	(68.70)	(60.36)	(23.94)
Latent Class Indicators				
LC1	---	0.553	0.553	13.199
		(16.10)	(16.10)	(17.45)
LC2	---	1.023	1.023	14.873
		(29.91)	(29.91)	(19.78)
LC3	---	0.543	0.543	12.913
		(15.38)	(15.38)	(17.01)
LC4	---	-0.493	-0.493	10.472
		(10.45)	(10.45)	(12.57)
LC6	---	1.165	1.165	13.875
		(32.40)	(32.40)	(18.21)
Age Effects				
Age	2.792	2.956	2.956	23.490
	(49.22)	(52.51)	(52.51)	(24.36)
Age Sq	-0.643	-0.678	-0.678	-7.540
	(50.88)	(54.14)	(54.13)	(24.55)
LC1 * Age	---	---	---	-19.811
				(20.44)
LC1 * Age-Squared	---	---	---	6.797
				(22.09)
LC2 * Age	---	---	---	-20.245
				(20.93)
LC2 * Age-Squared	---	---	---	6.768
				(22.01)
LC3 * Age	---	---	---	-17.959
				(18.43)
LC3 * Age-Squared	---	---	---	6.055
				(19.59)
LC4 * Age	---	---	---	-17.299
				(16.84)
LC4 * Age-Squared	---	---	---	6.060
				(19.06)
LC6 * Age	---	---	---	-20.091
				(20.67)
LC 6 * Age-Squared	---	---	---	6.902
				(22.41)
State Depen. Effects				
Arr_{t-1}	0.770	0.615	0.615	0.402
	(48.67)	(40.57)	(40.57)	(27.27)
Ancillary Parameters				
ln(r)	2.627	15.903	---	---
ln(s)	3.123	16.499	---	---
ln(alpha)	---	---	0.596	0.475
Log-Likelihood	-63700.517	-62157.225	-62157.225	-59971.658

Note: Absolute values of t-statistics are in parentheses. Latent class indicator variables are from the results presented in Chapter 6

2001), which is the nested model of the random effects negative binomial model.[7] This statistical test, which tests the significance of the r and s ancillary parameters, resulted in a χ^2_{01} value of 391.28 (p-value = 0.000), which clearly indicates the presence of significant unobserved heterogeneity. The estimate of the state dependence effect in the random effects model was 0.770 (t-statistic = 48.67), which is considerably larger in magnitude than the estimate from the semiparametric model. In fact, this estimate from the random effects model is nearly twice as large in magnitude. Although, it should be noted that both models are compatible in the sense that they each indicate a positive, significant, and substantively important effect of having been arrested at the prior age on the mean offending rate at the subsequent age. Still, there is a large discrepancy in the effect between the two models, which indicates a need for additional methods of calculating the estimate of this parameter in order to test its methodological robustness. In the third stage of this analysis, we use a method of calculating the state dependence parameter that has not previously been utilized in the empirical literature.

Stage Three—Incorporating Latent Class Indicator Variables

In the this stage, we employ the use of both the random effects negative binomial model and the standard negative binomial model, and estimate both of these models using the following specification:

[7] Similar to the discussion of testing the negative binomial model against the Poisson model, the null hypothesis here concerning the absence of individual-specific effects is also on the boundary of the parameter space. The null hypothesis, $\alpha_i=0$, is on the boundary of the parameter space because there can be an absence of individual-level effects, but they cannot be negative. Therefore, the appropriate statistical test is a boundary-value likelihood ratio test (Self and Liang 1987; Gutierrez et al. 2001). As shown by Self and Liang (1987), this test statistic is a 50:50 mixture (χ^2_{01}) of (1) a chi-square distribution with a point mass at zero (χ^2_{00}) and (2) a chi-square distribution with 1 degree of freedom (χ^2_1).

$$\ln(\lambda_{it}) = \beta_0 + \mathbf{X}_{LCV}\boldsymbol{\beta}_{LCV} + \left((age_{it}/10) * \beta_{age}\right) + \left(\left((age_{it}^2)/100\right) * \beta_{age^2}\right) + \left(arr_{t-1} * \beta_{arr_{t-1}}\right) + \varepsilon_{it}, \quad (3)$$

where $\mathbf{X}_{LCV}$ is a matrix of binary variables indicating latent class membership (with the adolescent-limited group, denoted as LC5 in Chapter 7, used as the reference group) and $\boldsymbol{\beta}_{LCV}$ is a column vector of regression coefficients for the latent class indicators.

The latent class indicator variables employed in this stage of the analysis are from the results of the semiparametric mixed Poisson models that were presented in Chapter 7. These latent class variables <u>*are not*</u> from the results presented in stage one of this chapter (i.e., Table 8.1). This stage of analysis allows us to address two previously unanswered questions raised earlier. First, does including the set of binary latent class indicator variables remove the underlying unobserved heterogeneity that was found in the random effects model just presented in stage two? Second, this stage of the analysis addresses whether the state dependence effects uncovered in stage two *over-estimates* the genuine state dependence effect because the age effects in the random effects model were controlled through the use of only two parameters that are assumed to be common to all individuals in the sample. The reader will recall that in Chapter 7 it was determined that allowing the age parameters to vary over the latent classes resulted in a significant improvement to the model fit. This indicates that all individuals in the sample do not follow the same trajectory of offending across the age distribution. Bushway et al. (1999) found that models that allowed for "time trend" or age effects significantly reduced the

effect of the state dependence variable, but the authors did not test to determine if it was also sensitive so as to allow the age parameters to vary over the latent classes. Presumably, the under-estimation of the age effects among the most frequent offenders will cause an over-estimation in the state dependence parameter because the state dependence parameter will then naturally absorb the "unaccounted for" age effects.

Model 2 of Table 8.2 presents the results of fitting a random effect negative binomial model with 5 latent class indicator variables included in the specification—the adolescent-limited group, denoted as LC5 in Chapter 7, was used as the omitted reference group. To test for the presence of significant unobserved heterogeneity in this model, we again performed a "boundary-value" likelihood ratio test that compares the log-likelihood value from this model against the log-likelihood value of the NB1 negative binomial model (which is negative binomial model with constant dispersion). This statistical test resulted in a χ^2_{01} value of 0.0 (p-value = 1.000), which indicates that the log-likelihood of the random effects model identically matched the log-likelihood from the NB1 negative binomial model, and more importantly, clearly indicates a lack of individual-level heterogeneity. For comparative purposes, we present the estimates from the NB1 negative binomial model in the Model 3 column of Table 8.2. In both of these models, one can see that (with the exception of the overall intercept term) all of the parameter estimates presented in Models 2 and 3 of Table 8.2 are identical, indicating that the individual-specific effects have been removed through the incorporation of the 5 latent class indicator variables. Importantly, note that the t-statistics for the latent class indicator variables are all highly significant, indicating that there are highly significant differences in the mean offending rates of the 5 latent classes and the omitted reference

group, the adolescent-limited group. With respect to the state dependence variable indicating an arrest at the prior age, the resulting parameter estimate in Model 3 (which is identical in Model 2 as well) was 0.615 (t-statistic = 40.57). Thus, the parameter estimate from the NB1 model with the latent class indicator variables is smaller than the estimate from the random effects model presented in Model 1 of Table 8.2, but it still is much larger than the estimate that resulted from the semiparametric model.

The NB1 model presented in Model 4 of Table 8.2 adds 10 parameters to the specification of the NB1 model presented in Model 3.[8] These additional parameters, which are interaction variables between each of the latent class indicator variables and the two age variables, allow each latent class to have its own set of regression coefficients for the age parameters. The overall age parameters (the "Age" and "Age- Squared" variables in Table 8.2) now represent the age parameters for the omitted reference group, the adolescent-limited (LC5) group. The latent class * age interaction variables in Model 4 represent the latent class-specific age coefficients (expressed in terms of deviations from the age coefficients of the omitted group, the adolescent-limited group). All of the latent class * age interaction variables were highly significant, indicating that that age coefficients for each of the latent class variables were significantly different from the age coefficients of the adolescent-limited group. The latent class indicator variables were also still highly significant in Model 4.

More importantly, however, the state dependence parameter estimate recovered in Model 4 is almost identical in magnitude to the parameter estimate from the

[8] The NB1 standard negative binomial model was employed here because the comparison of Models 2 and 3 indicated a complete lack of individual-specific effects once the latent class indicators are included in the specification. Since the NB1 model is a much simpler model in comparison with the random effects version, the NB1 model is used for Model 4.

semiparametric mixed Poisson model. The estimate in Model 4 of Table 8.2 was 0.402, whereas it was 0.404 in the 6-class model of Table 8.1. Thus, after accounting for the latent class-specific effects of the age coefficients, we were able to nearly recover the identical parameter estimates of the state dependence effect. Thus, had we simply applied the random effects estimator and not accounted for class-specific temporal shifts in the mean offending rate, we would have significantly over-estimated the magnitude of the state dependence parameter by almost twice its estimated value. As Bushway et al. (1999: 53) note, "clearly, then, it is possible for very general temporal shifts in the probability of offending activity to masquerade as genuine state dependence effects." We would further add that not only do general temporal shifts masquerade as state dependence effects, but so too do group-specific shifts.

Stage Four—Latent Class-Specific Models

In this analytical stage, we estimate separate NB1 negative binomial models on each latent class by itself.[9] The results presented up to this point have relied on a single estimate of the state dependence parameter, but for both theoretical and mathematical reasons, it is possible that the estimates calculated so far are not representative of the state dependence relationships *within each latent class itself.* Mathematically, the results calculated so far may simply be an "average" estimate that may reflect large effects in some latent classes and small/nonexistent effects in other latent classes. Recall that the

[9] These NB1 models were tested against the parametric random effects negative binomial models using the boundary-value likelihood ratio tests. In all 6 of the models, the null hypothesis of no individual-level effects could not be rejected. Similar to the results described in stage three, the resulting χ^2_{01} was equal to 0.00 in all 6 tests (p-value=1.00), which indicates that the log-likelihood values for the NB1 model (which ignores individual differences) were all identical to the log-likelihood values of the random effects models.

dual taxonomy theory of Moffitt (1993) specifically hypothesizes the state dependence effects vary over her two hypothetically discrete offender groups, with an especially pronounced effect predicted to occur among the adolescent-limited group and a muted or non-existent effect proposed among the life-course-persistent group that is dominated by a static etiological explanation. Sampson and Laub's (1993) theory, on the other hand, does not hypothesize that the state dependence process only affects specific types of offenders, but rather their general theory (i.e., same causal structure is at work for all individuals) specifies that the state dependence process should be found in all latent classes. According to Sampson and Laub, the state dependence process is unconditional, whereas it is a conditional process according to Moffitt. Of course, the effect does not causally exist in any latent class according to the theory of Gottfredson and Hirschi (1990).

By estimating models on each latent class separately, we are able to test for differential effects of the state dependence variable within each of the latent classes. The specification for this stage will be identical to the specification noted above in the second stage of analysis, and will include the following covariates: age, age-squared, and the binary indicator of arrest at the previous age. In other words, it is identical to the specification of equation (2), except here the standard negative binomial model is used and it is estimated for each latent class separately (see footnote 9). Again, the latent classes in Table 8.3 are from the analyses presented in Chapter 7, and are not the latent classes presented in Table 8.1 in this chapter. In Chapter 7, we presented results that indicated the fifth latent class, "LC5", in the 1981-82 sample clearly offended in an adolescent-limited fashion. Further, we also saw that "LC6" and "LC2" were the

Table 8.3. Investigation of State Dependence Effects with NB1 Negative Binomial Models: 1981-82 Sample, Latent Class-Specific Models

	Latent Class					
	LC1	LC2	LC3	LC4	LC5	LC6
Overall Intercept	-4.769	-2.983	-4.920	-7.454	-17.983	-3.879
	(36.79)	(28.90)	(31.77)	(19.67)	(22.04)	(23.71)
Age Effects						
Age	3.731	3.270	5.472	6.149	23.630	3.323
	(32.69)	(31.43)	(32.01)	(16.61)	(22.23)	(23.39)
Age-Squared	-0.754	-0.776	-1.471	-1.472	-7.586	-0.624
	(31.85)	(32.66)	(33.77)	(17.01)	(22.59)	(21.77)
State Depen. Effects						
Arr_{t-1}	0.383	0.365	0.433	0.508	0.399	0.437
	(13.58)	(13.47)	(12.71)	(6.08)	(6.13)	(11.79)
Ancillary Parameters						
ln(alpha)	0.247	0.597	0.466	0.174	0.414	0.691
Log-Likelihood	-15809.220	-17070.52	-11667.37	-3183.097	-3143.458	-9043.450
N (Panel)	16299	12123	12341	7675	5977	6038
N (Observations)	519	403	414	244	215	194

Note: Absolute values of t-statistics are in parentheses. The 6 latent classes are from the results presented in Chapter 6.

"highest rate" offenders, with an average of 44.10 and 36.35 arrest charges, respectively. LC4 had the lowest offending rate, and had, on average, a later age of onset (according to their arrest histories) in comparison with the other latent classes, but their offending did extend into the early thirties (albeit at a comparatively low-rate). The results of the latent class-specific models are presented in Table 8.3.

Overall, the results in Table 8.3 clearly speak to the robustness of the estimated state dependence effect calculated within each latent class in the 1981-82 sample. In all of the latent classes, the estimate is positive, significant, and substantively large. Furthermore, the estimates in all of the latent classes were generally in the immediate region of the overall parameter estimate (0.40) recovered in the 6-class model of Table 8.1 and Model 4 of Table 8.2. The largest effect estimated in any of the samples was in the fourth latent class, LC4. In this latent class, the estimated state dependence parameter was 0.508. The latent class with the lowest rate was LC2 (parameter estimate= 0.365; t-statistic=13.47).

Indeed, there is not even a hierarchical ranking of the magnitude of the state dependence parameter according to the rate or "style" of offending. The lowest rate latent class, LC4 had the largest parameter estimate, but the latent class with the highest rate of offending (LC6) had the second largest parameter estimate (0.437). The parameter estimate for the adolescent-limited group was nearly identical to the overall estimate presented in Tables 8.1 and 8.2, and was very close in magnitude to the effect calculated in the first latent class, LC1. A test of significant differences between each of the parameter estimates failed to generate any support for the hypothesis of differential

effects across the latent classes.[10] Instead, the overall estimate seems to accurately represent all of the latent classes in the 1981-82 sample. These results are compatible with the assertions of Sampson and Laub (1993), run counter to the expectations of Moffitt (1993), and do not support the eighth hypothesis (H_8) examined in this study. The adolescent-limited group had a state dependence parameter estimate that was both very similar to the overall estimate and to the estimates calculated in the other latent classes as well.

Stage Five—Post-Release Data Only

In the fifth and final analytical stage for the 1981-82 sample, we estimate models employing both the random effects negative binomial model and the standard negative binomial model, but here we only employ the post-release arrest data in generating the dependent variable of the models. We also include covariates identifying the theoretically relevant characteristics of wards (e.g., gang member, drug abuser) in the model. These results allow us to test the sensitivity of the results and ask the following two questions: (1) would the conclusions of this study have been any different had only the post-release arrest data been available for composing the dependent variable? (2) Are there any covariates significantly related to their post-release offending rate? Arguably, the datasets employed in this study are not typical of those commonly available to

[10] The method used here to test for significant differences between two coefficients is described in Clogg et al. (1995), Brame et al. (1998), and Paternoster et al. (1998). The test statistic is calculated as $z = \frac{\theta_1 - \theta_2}{\sqrt{(SE(\theta_1))^2 + (SE(\theta_2))^2}}$. Bonferonni adjustments were applied when calculating the p-values to adjust for the fact the multiple comparisons were being made. The same method of testing for significant differences between the latent-class specific state dependence estimates is employed in the analyses of both the 1986-87 and 1991-92 samples.

researchers in the field of criminology given the extended length of time over which these three samples are followed. Therefore, an interesting question concerns what, if any, differences would have resulted if we only had access to the post-release data. Furthermore, given that covariates were measured on the cases regarding their characteristics as of the time of the sample stay (i.e., before being released into the post-period), using the post-release data only allows us to include a set of covariates into the models to see what variables, net of control for unobserved heterogeneity, are predictive of their post-release offending rates. The analytical stage proceeds in two sections. First, we ignore the available covariates and estimate models like those presented in Table 8.2. These results pertaining to this section of the chapter are presented in Table 8.4. Second, we include the covariates in the model specifications. These results pertaining to this section of the chapter are presented in Table 8.5.

The results of four models are presented in Table 8.4. Model 1 of Table 8.4 is a standard NB1 negative binomial model that completely ignores individual differences in the propensity to offend. Other than the fact that this estimate is from a negative binomial model rather than a Poisson model, the state dependence parameter estimate presented in the first column of Table 8.4 is comparable to the 1-class parameter estimate presented in Table 8.1.[11] The estimate presented in Model 1 in Table 8.4 (0.847) is nearly identical in magnitude to the estimate presented in the 1-class model of Table 8.1

[11] The parameter estimate directly comparable to the 1-class model would be calculated using a standard Poisson model. We re-estimated Model 1 of Table 8.4 using a standard Poisson model. The corresponding estimate for the state dependence variable in the Poisson model was 0.864 (t-statistic = 69.37). Given the overdispersion in the data and the fact that neither the standard Poisson nor standard negative binomial models account for individual differences, these results were entirely expected. The parameter estimates should have been nearly identical in both models (i.e., neither controls for individual-specific effects), but

(0.857). Model 2 of Table 8.4 presents the estimates from a random effects negative binomial model applied to the post-release data of the 1981-82 sample. This model is directly comparable to Model 1 of Table 8.2. The state dependence parameter estimate presented in Model 2 (0.491) is much smaller in magnitude than the comparable parameter estimate found in Model 1 of Table 8.2 (0.770). However, it was still positive and highly significant (t-statistic = 24.15). Model 3 of Table 8.4 is directly comparable to Model 3 in Table 8.2.[12] The state dependence parameter estimate of the Model 3 in Table 8.4 was 0.415, whereas the comparable estimate using the full set of available data points was 0.615. Finally, Model 4 of Table 8.4 is directly comparable to Model 4 of Table 8.2, and once again we find that the parameter estimate using only the post-release data (0.316) was smaller in magnitude than the estimate using the full analytical dataset covering the entire available age distribution (0.402).

However, taken in their entirety, the results presented in Table 8.4 (using only the post-release data) produced findings strikingly similar to those making use of the full period of coverage. The state dependence parameter was still positive and highly significant, and furthermore, controlling for individual differences in criminal propensity significantly reduced the magnitude of the coefficient (e.g., compare Model 1 to Model 4).[13] These were the same results we observed when analyzing all of the available data.

---

the standard errors of the negative binomial model should have been larger (i.e., effects of accounting for the stochastic overdispersion).

[12] Estimating the comparable model to Model 2 of Table 8.2 produced a solution identical to Model 3 of Table 8.4. Again, the latent class indicators accounted for the unobserved heterogeneity.

[13] Of course, in reality, if one only had data on the post-release criminal activity of the cases, only Models 1 and 2 of Table 8.4 would be possible to estimate. The data used to derive the latent classes (in Chapter 7) that were used in Models 3 and 4 of Table 8.4 also included the pre-release data.

Table 8.4. Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1981-82 Sample, Post-Release Data Only

	Model			
	1	2	3	4
Overall Intercept	-1.252	-1.774	-3.894	-30.268
	(5.97)	(8.53)	(17.76)	(4.59)
Latent Class Indicators				
LC1	---	---	2.131	25.437
			(26.76)	(3.85)
LC2	---	---	2.478	26.968
			(31.11)	(4.09)
LC3	---	---	1.726	26.535
			(21.33)	(4.01)
LC4	---	---	0.567	27.142
			(6.00)	(4.03)
LC6	---	---	2.905	26.891
			(36.08)	(4.07)
Age Effects				
Age	0.975	1.324	1.561	31.966
	(6.16)	(8.41)	(10.07)	(4.86)
Age Sq	-0.271	-0.343	-0.384	-8.703
	(9.38)	(11.92)	(13.54)	(5.31)
LC1 * Age	---	---	---	-28.486
				(4.32)
LC1 * Age-Squared	---	---	---	8.041
				(4.91)
LC2 * Age	---	---	---	-28.746
				(4.36)
LC2 * Age-Squared	---	---	---	7.979
				(4.87)
LC3 * Age	---	---	---	-28.016
				(4.25)
LC3 * Age-Squared	---	---	---	7.636
				(4.65)
LC4 * Age	---	---	---	-30.182
				(4.53)
LC4 * Age-Squared	---	---	---	8.210
				(4.98)
LC6 * Age	---	---	---	-29.116
				(4.42)
LC 6 * Age-Squared	---	---	---	8.180
				(4.99)
State Depen. Effects	0.847	0.491	0.415	0.316
Arr_{t-1}	(46.42)	(24.15)	(22.48)	(17.23)
Ancillary Parameters				
ln(r)	---	1.309	---	---
ln(s)	---	1.293	---	---
ln(alpha)	0.666	---	0.462	0.411
Log-Likelihood	-39539.836	-38930.330	-37233.864	-36590.537
N (Panel)	37247	37247	37247	37247
N (Observations)	1989	1989	1989	1989

Note: Absolute values of t-statistics are in parentheses. Latent class indicator variables are from the results presented in Chapter 6.

Table 8.5 also presents 4 models, and they are directly comparable to the four models presented in Table 8.4, only here we add in the measured covariates into the specifications. The "Set of Control Variables" included in each of the models presented in Table 8.5 are the precise variables found in the models presented in Table 8.4. Thus, the control variables for Model 1 of Table 8.5 are age and age-squared, which are the two additional covariates found in Model 1 of Table 8.4.

Model 1 of Table 8.5 is a standard NB1 negative binomial that has no controls for unobserved individual differences. The only individual differences included in this model are the *measured* individual differences captured through the use of the covariates included in the model specification. Briefly, according to Model 1 of Table 8.5, the covariates that were found to be positively and significantly related to the arrest rate in the post-release period (found in bold type in Table 8.5) were African-American ethnicity, Hispanic ethnicity, sibling criminality, ineffective parental control, drug abuse, gang member, school dropout, juvenile court commitment, and the DDMS infraction rate while incarcerated in the CYA. Three variables were found to be significantly and negatively related to the arrest rate—first commitment, violent commitment offense, and commitment from Los Angeles County. Thus, it appears from Model 1 that there were a number of covariates significantly related to the arrest rate in the post-release period. Indeed, some of these variables such as drug abuse, school dropout, and gang membership could be interpreted as indicators of "state dependence" —e.g. being a school dropout increases the odds of criminal activity because it cuts off opportunities for conventional, prosocial activities. However, Gottfredson and Hirschi (1990) have argued

Table 8.5. Investigation of Subject Characteristic Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1981-82 Sample, Post-Release Data Only

	Model			
	1	2	3	4
Overall Intercept	-1.387	-1.943	-3.916	-30.096
	(6.48)	(9.00)	(17.47)	(4.57)
Set of Control Variables	A	A	B	C
Background/CYA Vars.				
African American	0.314	0.414	0.141	0.141
	(13.17)	(11.21)	(6.06)	(6.06)
Hispanic	0.121	0.165	0.086	0.090
	(4.42)	(3.92)	(3.21)	(3.40)
Other Ethnicity	-0.055	-0.158	0.028	0.033
	(0.77)	(1.52)	(0.40)	(0.48)
Family Violence	-0.038	-0.026	-0.014	-0.018
	(1.36)	(0.61)	(0.53)	(0.67)
Par. Alc./Drug.	-0.012	-0.020	-0.005	-0.003
	(0.53)	(0.55)	(0.21)	(0.15)
Par. Crim.	0.011	0.015	-0.026	-0.028
	(0.44)	(0.39)	(1.12)	(1.18)
Sibling Crim.	0.049	0.061	0.000	-0.002
	(2.61)	(2.08)	(0.00)	(0.09)
Neglect	-0.041	-0.034	0.004	0.004
	(1.86)	(1.00)	(0.17)	(0.17)
Control	0.043	0.068	0.033	0.037
	(2.01)	(2.01)	(1.60)	(1.79)
Abuse	0.026	0.015	0.009	0.013
	(0.92)	(0.33)	(0.32)	(0.48)
Sex. Abuse	-0.067	-0.117	-0.085	-0.095
	(0.68)	(0.82)	(0.89)	(0.99)
Drug Abuse	0.051	0.054	0.023	0.023
	(2.54)	(1.73)	(1.14)	(1.18)
Gang Member/Ass.	0.085	0.112	0.052	0.048
	(3.74)	(3.06)	(2.34)	(2.15)
School Dropout	0.067	0.092	-0.005	-0.007
	(3.52)	(3.09)	(0.29)	(0.36)
Juvenile Court	0.045	0.005	-0.026	-0.007
	(2.22)	(0.17)	(1.25)	(0.35)
First Commit.	-0.183	-0.207	-0.030	-0.030
	(8.49)	(5.99)	(1.43)	(1.40)
Comm. Off. Viol.	-0.121	-0.130	-0.035	-0.039
	(6.25)	(4.32)	(1.87)	(2.08)
Los Angles	-0.183	-0.025	-0.045	-0.043
	(8.91)	(0.76)	(2.20)	(2.12)
Infraction Rate	0.197	0.137	-0.006	-0.007
	(2.33)	(1.02)	(0.08)	(0.08)
State Depen. Effects				
Arr_{t-1}	0.795	0.476	0.406	0.307
	(43.08)	(23.45)	(21.88)	(16.65)
Ancillary Parameters				
ln(c)	---	1.370	---	---
ln(s)	---	1.388	---	---
ln(alpha)	0.648	---	0.461	0.409
Log-Likelihood	-29340.866	-28302.827	-37201.661	-36558.764

Notes: Absolute value of t-statistics in parentheses. Sets of control variables are as follows: Set A: Age, Age-Squared. Set B: Set A + Latent Class Indicators. Set C: Set A + Set B + entire set of interaction variables of the latent class indicators and the age variables.

that individuals self-select experiences that are consistent with their level of criminal propensity (i.e., self-control). Thus, individuals who have high criminal propensity also will likely drop out of school, join a gang, and abuse drugs. This argument is identical to the one explaining why there is an association between prior criminal offending and subsequent criminal offending. This association exists because the covariates are all simply proxy variables for a person's criminal propensity. In other words, to show that these variables have a *non-spurious* significant effect on the post-release arrest rate, we need to control for *unobserved* individual differences in criminal propensity. If Gottfredson and Hirschi are correct in their self-selection argument, these covariates should become non-significant after controlling for persistent individual differences.

Models 2 through 4 control for individual differences. Model 2 controls for individual differences through the use of the random effects estimator, whereas Models 3 and 4 (which are both NB1 negative binomial models) employ the latent class indicator variables. Given the assumption of the random effects model that all of the covariates included in the model are uncorrelated with the random effect (which is necessary in order to identify the random effects), Models 3 and 4 appear to offer the most compelling test of the significance of these covariates after controlling for individual differences. As indicated in both Models 3 and 4, most of the variables that were found to be significant in Model 1 were no longer found to be significant covariates. The only "state dependence" variable that maintained its statistical significance after controlling for individual differences in the propensity to offend was the gang member covariate. In other words, after controlling for individual differences in the propensity to offend, few variables were significantly related to the post-release arrest rates of these wards. This

conclusion is vastly different from what would have been obtained if we had used Models 1 or 2 of Table 8.5. Once again, controlling for persistent individual differences is critical for obtaining accurate estimates of the magnitude and significance of a given covariate.

With respect to the estimates of the state dependence variable found in Table 8.5, we find the results are virtually identical to those presented in Table 8.4. In other words, controlling for measured individual differences does not have any significant effect on the state dependence estimates.

RESULTS FOR THE 1986-87 SAMPLE

Attention is now turned to analysis of the data for the 1986-87 release sample. The five stage analytical approach undertaken here for this data set is identical to the approach presented above for the 1981-82 sample.

Stage One—The Semiparametric Mixed Poisson Model

We start again with a presentation of the results from the semiparametric mixed Poisson model. The same process described above for fitting and testing the semiparametric mixed Poisson model with varying numbers of "points of support" was also undertaken here. This model fitting process employed the use of both the BIC statistic and testing for global solutions.

The local solution testing that was completed for the 1986-87 sample indicated that both the 7-class and 8-class models were again prone to multiple local solutions that varied from one solution to another. Similar to the results for the 1981-82 sample, the 6-

class model generated the same unique solution all ten times this model was estimated. Furthermore, and as shown in Table 8.6, the 6-class model had the largest BIC value (i.e., least negative value), and thus the 6-class model was chosen as the model with the optimal number of latent classes. The 6-class model had a BIC value of –47723.16, whereas the BIC value for the 5-class model was only –47874.92. Again, since interest in this chapter also focuses on changes in the size of the state dependence coefficient, the solutions for all models up through the 6-class model are presented in Table 8.6.

Table 8.6 presents the results for the 1- through the 6-class semiparametric mixed Poisson models. The first column in Table 8.6 contains the 1-class semiparametric mixed Poisson model, which again is equivalent to the estimation of a standard Poisson regression model that makes no allowances for either stochastic variation or individual differences. The estimate of the state dependence parameter in the 1-class model that assumes complete homogeneity was 0.619.[14] This estimate was highly significant with a t-statistic calculated at a value of 51.58 (which supports hypothesis 5 of this study). The next column allows for two latent classes in the population, and here we find that this simple adjustment makes a large impact on the estimate of the state dependence parameter. In the 2-class model, the estimate of the state dependence parameter was reduced to 0.411 (t-statistic = 33.80). Allowing for another additional point of support, further reduced the state dependence parameter estimate to 0.345 (t-statistic = 27.98). Next, allowing for four points of supports reduced the estimate to 0.295 (t-statistic = 23.92), while allowing for five points of support reduced it to 0.266 (t-statistic = 21.37).

[14] The comparable state dependence estimate for the NB1 model was 0.645 (t-statistic = 35.51).

Table 8.6. Investigation of State Dependence Effects with Semiparametric Random Effects Poisson Model: 1986-87 Sample (N = 1443; Panel Observations =37390)

	Points of Support					
	1	2	3	4	5	6
Random Effects						
Class1	-3.953	-3.998	-7.785	-5.837	-6.922	-7.912
	(69.61)	(53.39)	(33.40)	(37.97)	(30.94)	(19.17)
Class 2	---	-6.860	-4.523	-4.701	-4.763	-4.971
		(38.95)	(39.54)	(32.35)	(26.00)	(22.73)
Class 3	---	---	-4.640	-4.416	-4.510	-2.987
			(34.30)	(31.89)	(33.90)	(6.15)
Class 4	---	---	---	-19.854	-23.4311	-23.5408
				(18.51)	(16.98)	(16.56)
Class 5	---	---	---	---	-4.535	-6.750
					(20.95)	(23.38)
Class 6	---	---	---	---	---	-4.115
						(11.30)
Age Effects						
Class 1 Age	3.883	5.945	9.009	6.405	7.983	8.9379
	(65.27)	(54.04)	(31.55)	(37.79)	(30.61)	(21.31)
Class 1 Age-Squared	-0.956	-0.902	-2.580	-1.716	-2.222	-2.4507
	(66.28)	(53.37)	(30.00)	(36.51)	(29.54)	(21.87)
Class 2 Age	---	7.881	5.2303	4.025	4.191	4.378
		(37.06)	(42.23)	(31.37)	(24.92)	(23.49)
Class 2 Age-Squared	---	-2.238	-1.3438	-0.817	-0.921	-0.9575
		(35.83)	(40.74)	(29.47)	(23.60)	(23.03)
Class 3 Age	---	---	3.986	5.177	5.326	3.8543
			(33.45)	(35.79)	(36.78)	(6.15)
Class 3 Age-Squared	---	---	-0.8199	-1.504	-1.361	-1.078
			(31.79)	(36.09)	(36.04)	(5.78)
Class 4 Age	---	---	---	25.708	30.515	30.8176
				(18.40)	(16.79)	(16.15)
Class 4 Age-Squared	---	---	---	-8.161	-9.753	-9.8903
				(18.08)	(16.43)	(15.59)
Class 5 Age	---	---	---	---	4.066	7.5113
					(21.50)	(24.82)
Class 5 Age-Squared	---	---	---	---	-0.799	-1.8336
					(19.82)	(23.93)
Class 6 Age	---	---	---	---	---	3.6824
						(12.44)
Class 6 Age-Squared	---	---	---	---	---	-0.7174
						(12.31)
State Depen. Effects						
Arr_{t-1}	0.619	0.411	0.345	0.295	0.266	0.245
	(51.58)	(33.80)	(27.98)	(23.92)	(21.37)	(19.18)
Log-Likelihood	-51270.407	-49327.309	-48628.149	-48057.750	-47802.170	-47635.868
BIC	-51284.956	-49256.407	-48671.795	-48115.946	-47874.915	-47723.162

Note: Absolute values of t-statistics are in parentheses.

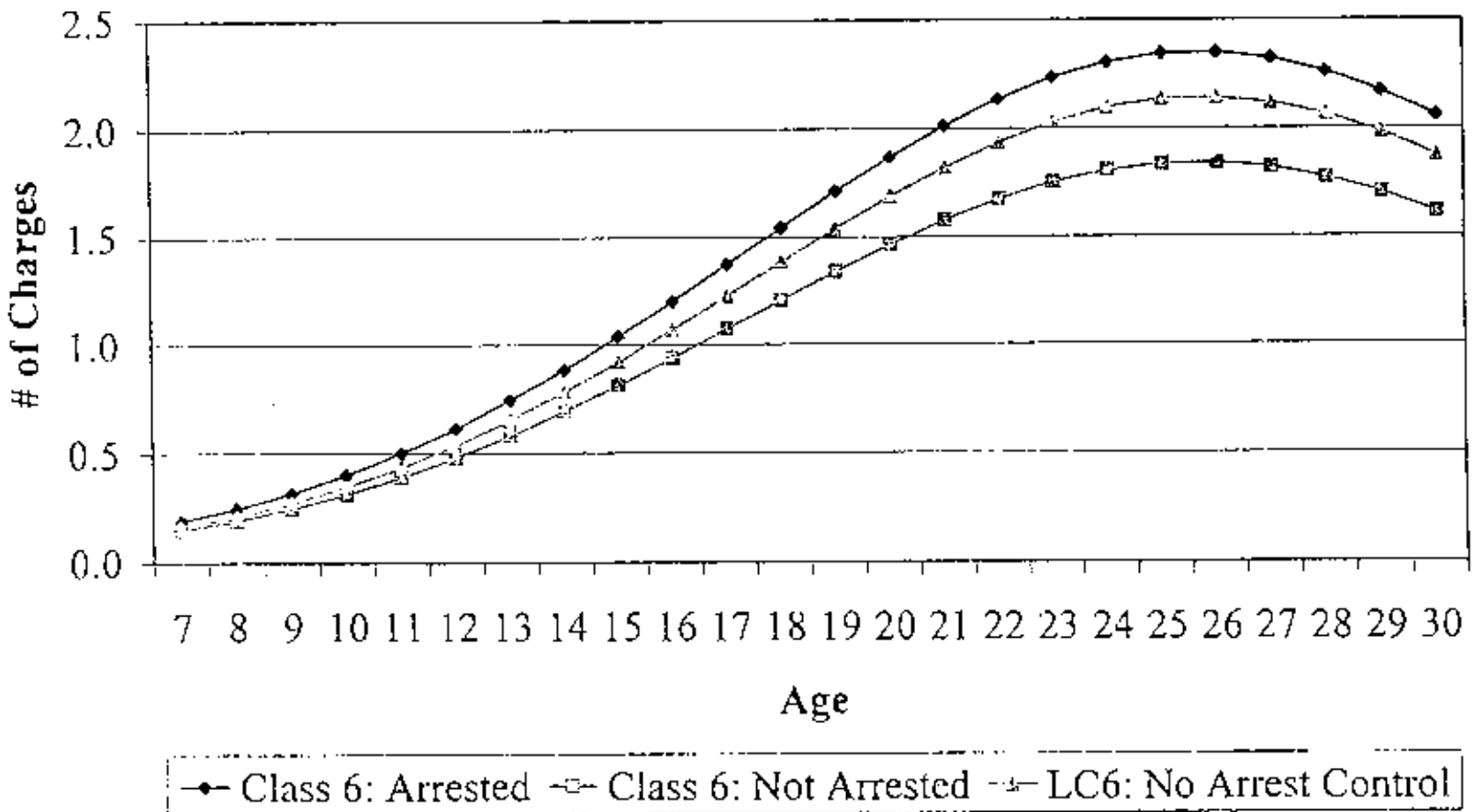
Finally, the sixth point of support reduced the state dependence parameter estimate to its final value of 0.245 (t-statistic = 19.18).[15]

Thus, accounting for the persistent individual differences in the 1986-87 sample by allowing for 6 points of support and class-specific age parameters reduced the magnitude of the state dependence parameter estimate from 0.619 (1-class model) to 0.245 (6-class model). This amounts to about a 60% reduction in the absolute size of the state dependence parameter. Yet, even after accounting for persistent individual differences, there was still a highly significant positive effect (which refutes the sixth hypothesis and supports our seventh hypothesis of this study).

To aid in the substantive interpretation of this parameter estimate, we predicted the arrest trajectory for the sixth latent class. This latent class is the same latent class labeled "LC6" in the 1986-87 results presented in Chapter 7—this group had the highest arrest charge total of any group and averaged over 36 total arrest charges. We have also included the predicted trajectory for this group that was generated in Chapter 7 with no control for arrest status at the prior age (labeled as "No Arrest Control" in Figure 8.2). Once again, the "No Arrest Control" trajectory (extracted from the results in Chapter 7) falls in between the estimated arrest rate for the cases that were not arrested at the prior age ("Not Arrested") and the estimated rate for the cases that were arrested at the prior age ("Arrested"). At the peak of their predicted offending rate, the cases that had been arrested at the prior age were estimated to be arrested for, on average, an extra 0.75 arrest charges. Thus, even though the size of the state dependence parameter estimate

[15] The estimates of the state dependence parameter in the local solution models of both the 7-class and 8-class models mixed Poisson model had the value of the parameter hovering at just above around 0.23

Figure 8.2. Predicted Mean Arrest Charges, by Arrest Status at Prior Age: 1986-87 Sample



decreased by over 60% after controlling for persistent individual differences in criminal propensity, the remaining parameter estimate still had a significant substantive implication for the arrest rates of the cases. Again, since these individuals are consistently arrested for serious offenses, we would argue that the state dependence parameter estimate in the 6-class model still implies a substantively important difference even though the magnitude of the effect was much smaller than was estimated in the model with no control for persistent individual differences.

Similar to the results generated for the 1981-82 sample, the results in the 1986-87 sample from the semiparametric mixed Poisson model clearly support the "mixed" position. There was significant population heterogeneity in the propensity to commit criminal acts, and accounting for these individual differences was important for calculating the magnitude of the state dependence parameter. Yet, even after controlling for the unobserved heterogeneity in the propensity to commit criminal acts, there was still a significant, positive, and substantively important relationship between the mean offending rate at a given age and whether an individual was arrested at the prior age. As shown in Figure 8.2 for one of the latent classes, individuals arrested at a prior age had a significantly higher mean arrest rate at the current age. Also, similar to the 1981-82 results, the addition of a second latent class to the model (or a second point of support) had the largest effect on the reduction of the magnitude of the state dependence effect (a decrease of 0.208), whereas the change that resulted from adding another point of support to the 5 point of support model was considerably smaller (a reduction of 0.021; i.e., 0.266-0.245 = 0.021).

Stage Two—The Parametric Random Effects Negative Binomial Model

In this stage we present the results from the parametric random effects negative binomial model and compare the estimate of the state dependence parameter in this model with the estimate from the semiparametric mixed Poisson model. Model 1 in Table 8.7 contains the resulting parameter estimates from the application of this model to the 1986-87 release sample. Before discussing the results with respect to the state dependence parameter estimate, we note that the boundary-value likelihood ratio test for the presence of significant unobserved heterogeneity (that tests the significance of the r and s ancillary parameters) resulted in a $\chi^2_{0:}$ value of 96.84 (p-value = 0.000). This finding indicates the presence of significant unobserved heterogeneity in the data. The estimate of the state dependence parameter in the parametric random effects model (with only a single set of age parameters) is found in the first column of Table 8.7. As reported in Table 8.7, the state dependence estimate in this model was 0.598, which was nearly identical in magnitude to the estimate generated in the 1-class semiparametric (Poisson) model that has no controls for persistent individual differences. Thus, as with the earlier sample, the random effects model for the 1986-87 data also significantly over-estimates the state dependence parameter in comparison with the final estimate we arrived at in the 6-class semiparametric model (which was 0.245). Comparing the estimate of the state dependence parameter in the parametric model to the estimate in the 6-class semiparametric model, we see the parametric model's estimate is nearly twice as large. This is a very large discrepancy between the estimates of these two models.

Table 8.7. Investigation of State Dependence Effects with Parametric Random Effects & NB1 Negative Binomial Models: 1986-87 Sample

	Model			
	1	2	3	4
Overall Intercept	-4.755	-5.726	-5.091	-21.033
	(57.64)	(62.91)	(56.07)	(20.93)
Latent Class Indicators				
LC1	---	0.534	0.534	13.375
		(12.82)	(12.82)	(13.03)
LC2	---	0.753	0.753	16.131
		(18.20)	(18.20)	(15.83)
LC3	---	1.066	1.066	17.938
		(25.15)	(25.15)	(17.65)
LC5	---	1.273	1.273	14.420
		(29.70)	(29.70)	(14.01)
LC6	---	1.247	1.247	17.022
		(28.75)	(28.75)	(16.60)
Age Effects				
Age	4.014	4.246	4.246	27.101
	(46.32)	(49.46)	(49.46)	(21.17)
Age Sq	-0.983	-1.037	-1.037	-8.571
	(46.62)	(49.76)	(49.76)	(21.30)
LC1 * Age	---	---	---	-18.663
				(14.37)
LC1 * Age-Squared	---	---	---	6.310
				(15.52)
LC2 * Age	---	---	---	-22.798
				(17.68)
LC2 * Age-Squared	---	---	---	7.636
				(18.91)
LC3 * Age	---	---	---	-23.193
				(17.98)
LC3 * Age-Squared	---	---	---	7.495
				(18.53)
LC5* Age	---	---	---	-19.914
				(15.34)
LC5 * Age-Squared	---	---	---	6.845
				(16.88)
LC6 * Age	---	---	---	-23.534
				(18.20)
LC 6 * Age-Squared	---	---	---	7.877
				(19.48)
State Depen. Effects				
Arr_{t-1}	0.598	0.465	0.465	0.248
	(31.71)	(25.29)	(25.29)	(13.90)
Ancillary Parameters				
ln(r)	3.254	15.742	---	---
ln(s)	3.870	16.377	---	---
ln(alpha)	---	---	0.635	0.627
Log-Likelihood	-41852.752	-41035.307	-41035.306	-41615.098

Note: Absolute values of t-statistics are in parentheses. Latent class indicator variables are from the results presented in Chapter 6

Stage Three—Incorporating Latent Class Indicator Variables

In the this stage, we again employ the use of the random effects and standard (NB1) negative binomial models, and build on the model specification used in Model 1 of Table 8.7. Model 2 in Table 8.7 is a parametric random effects model that also includes a set of binary latent class indicators. Model 3 is the NB1 negative binomial model (that does not include individual specific effects as in Model 2). The latent class indicator variables employed in Models 2 through 4 of this stage are from the results of the semiparametric mixed Poisson models presented in Chapter 7. The adolescent-limited group, denoted as LC4 in the Chapter 7 analyses of the 1986-87 data, was used as the omitted reference group in these models.

Comparing the results of Models 2 and 3 in Table 8.7, we find that the two models generate identical parameter estimates and solutions. The boundary-value likelihood ratio test comparing these two models generates a χ^2_{01} value of 0.001 (p-value = 1.000), which indicates that including just the set of five binary indicators removed the presence of significant unobserved heterogeneity (which was previously found to be significant in stage two). The state dependence parameter estimate in Model 3 was 0.465, which is smaller than the estimate from the parametric random effects model presented in Model 1 of Table 8.7, but is still much larger than the estimate that we arrived at in the 6-class finite mixture model (0.245). Thus, while the indicator variables do remove the presence of significant individual-level heterogeneity, the indicator variables by themselves do not allow for the recovery of the same parameter estimate found in the 6-class mixture model.

In Model 4 of Table 8.7, we next added a set of interaction variables between the latent class indicator variables and the age and age-squared variables to the model. In this part of the analysis, we are interested in determining whether the state dependence effects uncovered in stage two is *over-estimated* because the age effects in the random effects model (Model 2) were controlled through the use of only two parameters assumed to be common to all individuals in the sample. In the analyses of the 1986-87 data found in Chapter 7, it was determined that allowing the age parameters to vary over the latent classes resulted in a highly significant improvement to the model fit. This was an indication that all individuals in the 1986-87 sample do not follow the same trajectory of offending across the age distribution. Our presumption is that the under-estimation of the age effects among the most frequent offenders will cause an over-estimation in the state dependence parameter. We believe this will occur because this parameter will absorb the unaccounted for "age effects" in the most "active" latent classes.

The NB1 model presented in Model 4 of Table 8.7 adds 10 parameters to the specification of the NB1 model found in Model 3. These additional parameters allow each latent class to have its own set of regression coefficients for the age parameters. Thus, if the overestimation of the state dependence parameter is a consequence of erroneously modeling the age effects with only two parameter estimates, we should find a state dependence estimate in Model 4 that is similar to the point estimate found in the 6-class semiparametric model. The parameter estimates for this model are found in the fourth numerical column of Table 8.7. First, note that all of the latent class * age interaction variables in Model 4 were highly significant, which indicates that the offending trajectories of these 5 latent classes were significantly different from the

offending trajectory of the adolescent-limited group. Again, all of the latent class indicator variables were also still highly significant in Model 4.

More important for our concerns here though, allowing for latent class-specific age parameters resulted in the successful recovery of essentially the exact same state dependence parameter estimate found in the semiparametric mixed Poisson model in Table 8.6. The state dependence parameter estimate in Model 4 of Table 8.7 was 0.248, whereas the estimate found in the 6-class model of Table 8.6 was 0.245. Identical to the previous findings in the 1981-82 sample, we find once again that the calculation of the estimate of the state dependence parameter in the 1986-87 sample is very sensitive to the accurate estimation of the "age effects" in these data. Clearly, unaccounted for age effects in the data masquerade as indicators of genuine state dependence.

Stage Four—Latent Class-Specific Models

Next, We present the results from the models estimated on each latent class by themselves. Again, the purpose of estimating these models is to examine whether the overall estimate of the state dependence effect is reflective of the state dependence relationship within a given latent class. Recall, that one of the main theoretical reasons for estimating the state dependence effect within the latent classes was to determine if (1) the adolescent-limited group had a much larger state dependence effect and (2) if there was a minimal state dependence effect in the "life-course persistent" group. In Chapter 7, the adolescent-limited group was the fourth latent class, and hence it was labeled "LC4." The two classes with the highest arrest rates were the fifth and sixth latent classes (averaging 36.2 and 35.9 arrest charges respectively). The fifth latent class had a higher

and earlier peak rate of offending (3 arrest charges at about age 20), whereas the sixth latent class did not reach their peak age of offending until their late twenties (where it peaked at about 2 arrest charges). The results of estimating NB1 negative binomial models on each of the latent classes separately are found in Table 8.8.[16]

As depicted in Table 8.8, the estimated state dependence effects were positive in all six of the latent classes, however, in the fifth latent class, the estimate was not statistically significant at the 0.05 level (t-statistic = 1.81; p-value = 0.070). This latent class represented just under 12% of the sample, and thus a significant, positive state dependence effect was estimated in the latent classes representing over 88% of the 1986-87 release sample. For the first three latent classes, the estimated state dependence relationships ranged from 0.229 to 0.269 and were very similar to the overall effect of 0.245. For the fourth and sixth latent classes, the estimated state dependence relationships were of a slightly greater magnitude (both estimated at just over 0.38). The 95% confidence intervals for the state dependence parameter estimates of these two latent classes, however, both overlapped with the 95% confidence intervals for the first three latent classes. Indeed, a test of significant differences between the parameter estimates of the two latent classes with the largest effects and the parameter estimates of the first three groups (which had smaller estimates) failed to reject the null hypothesis indicating no differences. In other words, taking into account sampling variation, the hypothesis that the five significant state dependence coefficients were not significantly different from

[16] In an identical finding to the 1981-82 sample, the boundary-value likelihood ratio tests of these NB1 models against the parametric random effect models (which allow for individual-level effects) resulted in a failure to reject the null hypothesis of no individual-level effects for all six latent classes. The value of χ^2_{01} in all six of the tests was equal to 0.00 (p-value=1.00), which indicates that the log-likelihood values for the NB1 model (which ignores individual differences) were all identical to the log-likelihood values of the random effects models

Table 8.8. Investigation of State Dependence Effects with NB1 Negative Binomial Models: 1986-87 Sample, Latent Class-Specific Models

	Latent Class					
	LC1	LC2	LC3	LC4	LC5	LC6
Overall Intercept	-7.803	-4.951	-3.096	-20.238	-6.717	-3.836
	(33.93)	(28.07)	(18.62)	(18.94)	(27.13)	(17.86)
Age Effects						
Age	8.575	4.311	3.918	26.046	7.415	3.394
	(33.58)	(25.55)	(20.50)	(19.01)	(29.24)	(16.99)
Age-Squared	-2.300	-0.938	-1.078	-8.253	-1.774	-0.664
	(34.26)	(24.57)	(21.61)	(19.26)	(29.19)	(15.22)
State Depen. Effects						
Arr_{t-1}	0.229	0.269	0.239	0.386	0.080	0.388
	(5.95)	(7.54)	(5.76)	(5.06)	(1.81)	(8.24)
Ancillary Parameters						
ln(alpha)	0.288	0.341	0.518	0.555	0.713	0.699
Log-Likelihood	-8708.972	-9389.148	-7153.637	-2375.699	-6178.243	-5611.288
N (Panel)	10207	9068	5326	4610	4336	3843
N (Observations)	398	333	211	188	170	143

Note: Absolute values of t-statistics are in parentheses. Latent class indicator variables are from the results presented in Chapter 6.

one another could not be rejected. Taken as a whole, and with the exception of the one latent class with a non-significant positive effect, the overall state dependence effect seemed to accurately represent the vast majority of the latent classes in the sample. Thus, the results for the 1986-87 sample also refute Moffitt's hypothesis that the state dependence effects are much greater in the adolescent peaked (or adolescent-limited) offender group relative to the life-course-persistent group (hypothesis eight of this study).

Stage Five—Post-Release Data Only

In the last stage of analysis for the 1986-87 sample, we limit the analyses to using dependent variables constructed from only the data in the post-release period and include covariates indicating the background characteristics of the cases to see if any of the covariates are predictive of the post-release arrest rate. Limiting the analyses to dependent variables compiled only from the post-release data allows for sensitivity analyses to determine whether the conclusions of the study would have changed had the available data only covered a much more limited age range.

The results of the sensitivity analyses are presented in first four models of Table 8.9. Model 1 of Table 8.9 is a standard NB1 negative binomial model that completely ignores individual differences in the propensity to offend. The state dependence parameter estimate presented in the first column of Table 8.9 is comparable to the 1-class parameter estimate presented in Table 8.6, and we see here that the estimate is of a similar magnitude to the one presented in Table 8.6 (0.663 versus 0.619).[17] The next

[17] The parameter estimate directly comparable to the 1-class model would be calculated using a standard Poisson model. We re-estimated Model 1 of Table 8.9 using a standard Poisson model. The corresponding estimate for the state dependence variable in the Poisson model was 0.678 (t-statistic = 42.25).

Table 8.9. Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1986-87 Sample, Post-Release Data Only

	Model			
	1	2	3	4
Overall Intercept	-2.343	-3.173	-5.566	-23.335
	(6.18)	(8.39)	(14.45)	(3.12)
Latent Class Indicators				
LC1	---	---	1.774	15.186
			(17.78)	(2.01)
LC2	---	---	2.357	15.838
			(23.81)	(2.11)
LC3	---	---	2.333	19.899
			(23.25)	(2.64)
LC5	---	---	2.883	14.756
			(28.84)	(1.96)
LC6	---	---	3.025	13.697
			(30.23)	(1.82)
Age Effects				
Age	2.131	2.622	3.086	24.458
	(7.06)	(8.67)	(10.46)	(3.35)
Age Sq	-0.535	-0.648	-0.739	-6.739
	(9.11)	(11.00)	(12.77)	(3.80)
LC1 * Age	---	---	---	-16.336
				(2.22)
LC1 * Age-Squared	---	---	---	4.689
				(2.63)
LC2 * Age	---	---	---	-18.656
				(2.55)
LC2 * Age-Squared	---	---	---	5.604
				(3.15)
LC3 * Age	---	---	---	-20.859
				(2.84)
LC3 * Age-Squared	---	---	---	5.839
				(3.28)
LC5* Age	---	---	---	-15.882
				(2.16)
LC5 * Age-Squared	---	---	---	4.784
				(2.69)
LC6 * Age	---	---	---	-16.908
				(2.31)
LC 6 * Age-Squared	---	---	---	5.360
				(3.02)
State Depen. Effects				
Arr_{t-1}	0.663	0.416	0.293	0.144
	(27.56)	(15.49)	(12.13)	(5.97)
Ancillary Parameters				
ln(r)	---	1.554	---	---
ln(s)	---	1.767	---	---
ln(alpha)	0.746	---	0.527	0.428
Log-Likelihood	-22953.843	-22733.245	-21657.142	-21107.500
N (Panel)	20090	20090	20090	20090
N (Observations)	1443	1443	1443	1443

Note: Absolute values of t-statistics are in parentheses. Latent class indicator variables are from the results presented in Chapter 6.

model in Table 8.9 (denoted as Model 2) presents the estimates from a random effects negative binomial model applied to the post-release data of the 1986-87 sample. This model is directly comparable to Model 1 of Table 8.7. The state dependence parameter estimate in Model 2 here was 0.416, which is smaller than the estimate found in Model 1 of Table 8.7 (0.598). However, the estimate from Model 2 in Table 8.9 was still positive and highly significant (t-statistic = 15.49), which is consistent with the results presented earlier in Table 8.7 where the entire set of available data points was used to construct the dependent variable.

The next model depicted in Table 8.9, Model 3, is directly comparable to Model 3 of Table 8.7, and the state dependence parameter estimate in this model was estimated at 0.293.[18] The comparable estimate from the model making full use of all available data points was 0.465. Note that we are consistently finding a marginally smaller estimate of the state dependence estimate when we only use the more limited time period covered in the post-release data. Finally, the last model found in Table 8.9 is directly comparable to Model 4 of Table 8.7. Again we find that the parameter estimate using only the post-release data (estimate = 0.144) was smaller in magnitude than was the estimate using the full analytical dataset covering the entire available age distribution (0.245). However, it was still positive and significantly related to the mean rate of offending, even if the relationship was not as strong with this more limited set of data.

When viewed in their entirety, thus far, the results depicted in Table 8.7 (using only the post-release data) have produced findings consistent with those we arrived at

[18] Estimating the comparable model to Model 2 of Table 8.7 (i.e., a random effects model with the latent class indicators) produced a solution identical to Model 3 of Table 8.9.

when making use of data covering a more extended period of time. The state dependence parameter in the last column of Table 8.7 was still positive and highly significant. Furthermore, controlling for individual differences in criminal propensity and differences in the age parameters significantly reduced the magnitude of the state dependence coefficient (e.g., 0.663 in Model 1 to 0.144 in Model 4).[19]

The last set of results for the 1986-87 sample are presented in Table 8.10. In these models we simply add the background characteristics to the specifications used in the models found in Table 8.9. Interest in this table concerns whether any covariates are significantly related to the post-release arrest rate, especially after controlling for unobserved heterogeneity. Model 1 of Table 8.10 is the "naïve" model, which assumes that there are no individual differences in the propensity to commit criminal acts. In this model, the following covariates were significantly related to the post-release arrest rate (they are in bold type in Table 8.10): African American ethnicity, Hispanic ethnicity, parental alcohol/drug dependence, drug abuse, school dropout, juvenile court commitment, first commitment, coming from Los Angeles County, and the DDMS infraction rate. Model 2, which is a random effects negative binomial model, essentially leaves the results from Model 1 unchanged. The random effects model assumes that the random effects (which are mathematically pulled out of the likelihood expression) are uncorrelated with the included covariates, a scenario that is inconsistent with the theory of Gottfredson and Hirschi (i.e., those with high criminal propensity should be more likely to be drug abusers, gang members, and school dropouts). As shown by Brame et

[19] Of course, if one only had data on the post-release criminal activity of the cases, we could not have estimated Model 3 or Model 4 of Table 8.9 because the pre-release arrest data was used to derive the latent classes (in Chapter 7).

Table 8.10. Investigation of Subject Characteristic Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1986-87 Sample, Post-Release Data Only

	Model			
	1	2	3	4
Overall Intercept	-2.175	-3.040	-5.340	-23.421
	(5.70)	(7.92)	(13.69)	(3.14)
Set of Control Variables	A	A	B	C
Background/CYA Vars.				
African American	**0.241**	**0.290**	**0.114**	**0.106**
	(6.99)	**(6.29)**	**(3.43)**	**(3.23)**
Hispanic	**0.132**	**0.152**	0.063	0.058
	(3.78)	**(3.25)**	(1.83)	(1.73)
Other Ethnicity	0.056	0.011	-0.021	-0.054
	(0.67)	(0.10)	(0.26)	(0.68)
Family Violence	-0.052	-0.052	-0.048	-0.043
	(1.41)	(1.05)	(1.35)	(1.24)
Par. Alc./Drug.	**0.068**	0.064	0.013	0.016
	(2.42)	(1.67)	(0.46)	(0.59)
Par. Crim.	0.024	0.044	0.016	0.022
	(0.85)	(1.15)	(0.59)	(0.80)
Sibling Crim	0.018	0.023	0.011	0.024
	(0.74)	(0.69)	(0.46)	(1.02)
Neglect	0.041	0.052	0.024	0.030
	(1.54)	(1.46)	(0.91)	(1.18)
Control	0.000	-0.007	-0.018	-0.017
	(0.01)	(0.17)	(0.60)	(0.58)
Abuse	-0.061	-0.051	0.015	0.015
	(1.74)	(1.08)	(0.43)	(0.45)
Sex. Abuse	-0.069	-0.127	-0.079	-0.023
	(0.91)	(1.27)	(1.08)	(0.31)
Drug Abuse	**0.202**	**0.230**	**0.051**	0.043
	(5.33)	**(4.65)**	(1.37)	(1.17)
Gang Member/Ass	-0.027	-0.000	-0.007	-0.004
	(0.96)	(0.00)	(0.26)	(0.15)
School Dropout	**0.061**	**0.074**	0.035	0.015
	(2.48)	**(2.20)**	(1.47)	(0.63)
Juvenile Court	**-0.101**	**-0.128**	-0.046	0.052
	(3.31)	(3.12)	(1.46)	(1.63)
First Commit.	**-0.250**	**-0.288**	**-0.125**	**-0.097**
	(9.98)	**(8.41)**	**(5.00)**	**(3.90)**
Comm. Off. Viol.	-0.031	-0.036	-0.037	-0.034
	(1.19)	(1.02)	(1.44)	(1.35)
Los Angles	**-0.186**	**-0.116**	0.045	0.020
	(6.89)	**(3.15)**	(1.68)	(0.74)
Infraction Rate	**0.266**	**0.269**	**0.239**	**0.218**
	(3.04)	**(2.16)**	(2.70)	**(2.49)**
State Depen. Effects				
Arr_{t-1}	**0.622**	**0.414**	**0.285**	**0.135**
	(25.74)	**(15.44)**	**(11.78)**	**(5.60)**
Ancillary Parameters				
ln(r)	---	1.708	---	---
ln(s)	---	1.969	---	---
ln(alpha)	0.722	---	0.525	0.427
Log-Likelihood	-22808.490	-22634.446	-21622.045	-21083.646

Notes: Absolute value of t-statistics in parentheses. Sets of control variables are as follows: Set A: Age, Age-Squared. Set B: Set A + Latent Class Indicators. Set C: Set A + Set B + entire set of interaction variables of the latent class indicators and the age variables.

al. (1999), if the random effects are correlated with the included covariates, the parameter estimates of the covariates will be positively biased (and favor a rejection of the null hypothesis). Thus, a better test makes use of the latent class indicators to control for the unobserved heterogeneity. Models 3 and 4 include such indicators in the specification, and similar to the results obtained for the 1981-82 sample, we see that most of the significant covariates in Model 1 (and Model 2) now become insignificant. In fact, in these two models (Models 3 and 4), there were only 4 covariates found to significantly relate to the post-release arrest rate: African American ethnicity, first commitment, DDMS infraction rate, and the binary indicator of arrest at the immediately prior age. The state dependence estimate in Model 4 was essentially the same as it was in Model 4 of Table 8.9, and thus including background characteristics in the equation did not significantly change the estimate of the state dependence relationship.

RESULTS FOR THE 1991-92 SAMPLE

Stage One—The Semiparametric Mixed Poisson Model

The results for the 1991-92 sample begin again with the estimation of the semiparametric mixed Poisson model. The same process followed in the first two samples for fitting and testing the semiparametric mixed Poisson model with varying numbers of "points of support" was also employed here with the 1991-92 sample. The BIC statistic and testing for global solutions were again used to find the optimal number of components in the mixing distribution. Consistent with both the 1981-82 and 1986-87 samples, the local solution testing indicated that both the 7-class and 8-class models were

prone to multiple local solutions. The 6-class model in this sample also generated the same unique solution all ten times this model was estimated. As shown in Table 8.11, the BIC statistic also favored the 6-class model (BIC = -36533.44) over the 5-class model (-36691.35). However, since it is important to see the changes in the size of the state dependence coefficient as the heterogeneity distributed is better approximated with dditional points of support, the solutions for all models up through the 6-class model are presented in Table 8.11.

Results for the 1- through 6-class semiparametric mixed Poisson models are presented in Table 8.11. The first column in Table 8.11 contains the 1-class Poisson model. The estimate of the state dependence parameter from the 1-class model (that assumes complete homogeneity) was 0.553, which was highly significant with a t-statistic of 40.25 (offering support for hypothesis 1 of this study).[20] Adding an additional point of support (i.e., two latent classes), we again find a large impact on the estimate of the state dependence parameter. In the 2-class model, the estimate of the state dependence parameter was reduced to 0.375 (t-statistic = 27.00). Allowing for 3 points of support reduced the state dependence parameter estimate to 0.319 (t-statistic = 22.79), while adding another point of support (4-class model) reduced the estimate to 0.262 (t-statistic = 18.51). With five points of support, the state dependence parameter estimate was 0.229 (t-statistic = 15.71), and finally, the final estimate from the six point of support model was 0.206 (t-statistic = 14.38).[21]

[20] The comparable state dependence estimate for the NB1 model was 0.549 (t-statistic = 26.44).
[21] The estimates of the state dependence parameter in the local solution models of both the 7-class and 8-class models mixed Poisson model had the value of the parameter at just over 0.20

Table 8.11. *Investigation of State Dependence Effects with Semiparametric Random Effects Poisson Model: 1991-92 Sample (N = 1434; Panel Observations = 29385)*

	Points of Support					
	1	2	3	4	5	6
Random Effects						
Class 1	-6.004	-10.5233	-7.454	-7.978	-10.749	-7.177
	(70.12)	(32.62)	(37.13)	(27.71)	(4.46)	(21.78)
Class 2	---	-5.5445	-5.582	-7.303	-5.149	-17.245
		(42.14)	(32.51)	(26.04)	(15.19)	(14.28)
Class 3	---	---	-25.906	-5.741	-7.172	-7.877
			(19.24)	(33.94)	(17.73)	(29.46)
Class 4	---	---	---	-25.443	-30.308	-5.701
				(20.37)	(5.78)	(15.03)
Class 5	---	---	---	---	-6.070	-38.839
					(5.24)	(15.74)
Class 6	---	---	---	---	---	-4.793
						(15.59)
Age Effects						
Class 1 Age	6.5983	12.8995	8.695	8.269	12.070	7.353
	(66.64)	(32.82)	(34.75)	(23.80)	(3.72)	(18.54)
Class 1 Age-Squared	-1.7859	-3.8711	-2.465	-2.142	-3.321	-1.867
	(66.29)	(31.07)	(32.30)	(21.31)	(3.29)	(16.24)
Class 2 Age	---	5.9493	5.970	9.349	4.758	20.636
		(42.30)	(33.90)	(28.16)	(14.05)	(13.92)
Class 2 Age-Squared	---	-1.494	-1.449	-2.786	-1.049	-6.011
		(41.41)	(32.61)	(27.52)	(9.68)	(13.31)
Class 3 Age	---	---	33.783	6.166	8.362	9.264
			(18.56)	(34.81)	(16.12)	(31.00)
Class 3 Age-Squared	---	---	-10.7928	-1.482	-2.174	-2.445
			(17.69)	(32.17)	(13.96)	(29.88)
Class 4 Age	---	---	---	32.970	39.928	7.780
				(19.71)	(5.09)	(17.25)
Class 4 Age-Squared	---	---	---	-10.453	-12.872	-2.445
				(18.83)	(4.51)	(18.64)
Class 5 Age	---	---	---	---	8.261	52.794
					(5.83)	(15.54)
Class 5 Age-Squared	---	---	---	---	-2.596	-17.584
					(6.27)	(15.12)
Class 6 Age	---	---	---	---	---	4.586
						(14.61)
Class 6 Age-Squared	---	---	---	---	---	-0.976
						(12.63)
State Depen. Effects						
Arr_{it}	0.553	0.375	0.319	0.262	0.229	0.206
	(40.25)	(27.00)	(22.79)	(18.51)	(15.71)	(14.38)
Log-Likelihood	-39101.26	-37583.21	-37113.93	-36811.490	-36618.66	-36446.22
BIC	-39115.8	-37612.28	-37157.54	-36869.64	-36691.35	-36533.44

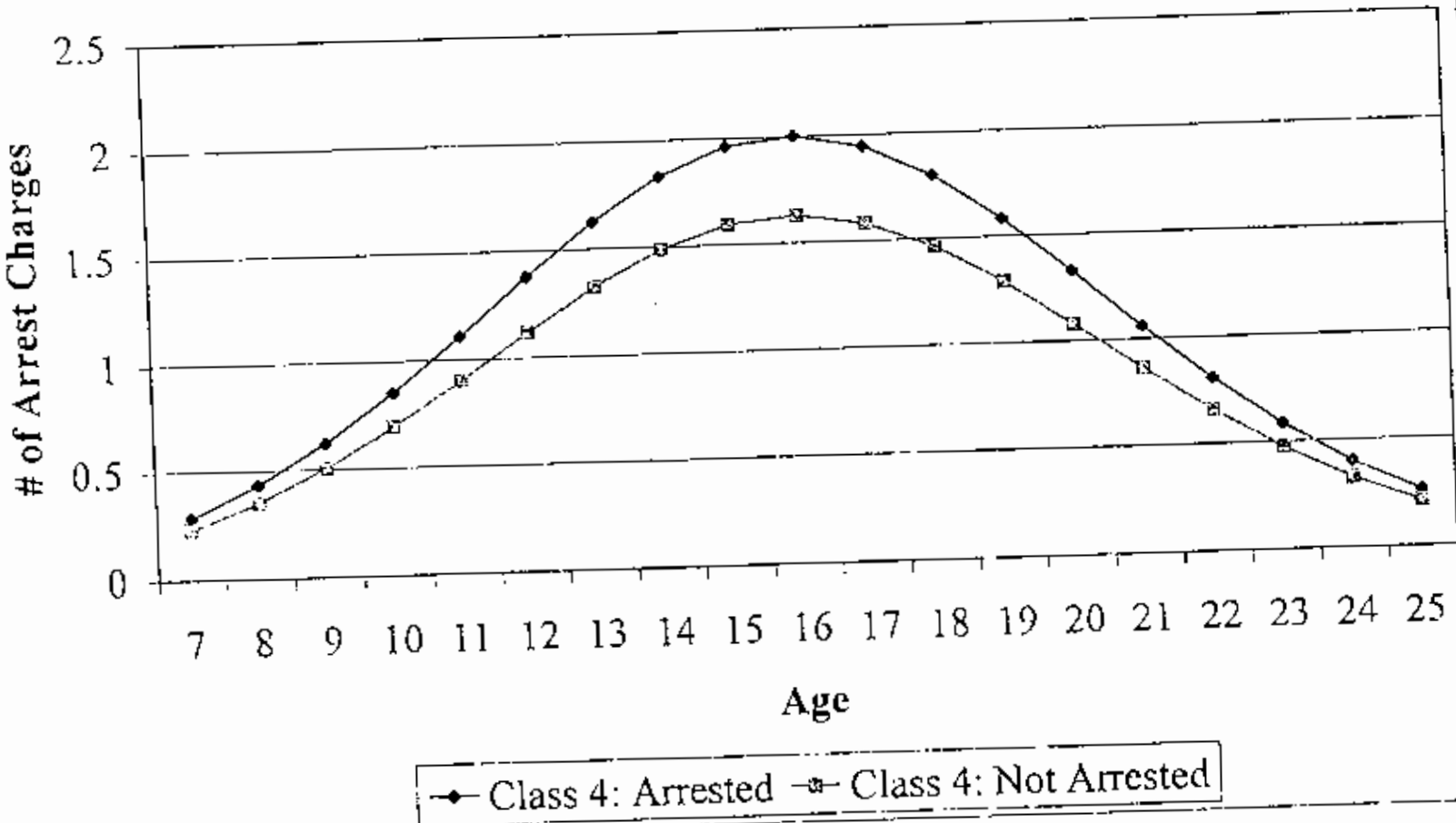
Note: Absolute values of t-statistics are in parentheses.

Similar to the two previous samples, accounting for the persistent individual differences in the 1991-92 sample significantly reduced the magnitude of the state dependence parameter estimate from 0.553 (1-class model) to 0.206 (6-class model), which is a 63% reduction in the absolute size of the state dependence parameter. But even after accounting for the unobserved heterogeneity (e.g., persistent individual differences), the state dependence parameter was still highly significant (which refutes hypothesis 6 and supports hypothesis 7 of this study).

Similar to the analyses of the two earlier samples, we also predicted an arrest trajectory for the fourth latent class in Table 8.11 to show the substantive implications of the state dependence estimate. This latent class corresponds to the "LC4" latent class presented in the results section for the 1991-92 sample in Chapter 7. As can be seen in Figure 8.3, the estimated arrest rate at a given age for the cases that were not arrested at the prior age ("Not Arrested") were predicted to be significantly lower than the estimated rate for the cases that were arrested at the prior age ("Arrested"). At the peak of their predicted offending rate, the cases that had been arrested at the prior age were estimated to be arrested for, on average, an additional 0.50 arrest charges. Thus, in spite of the fact that there was a significant decrease in the state dependence parameter estimate after controlling for persistent individual differences in criminal propensity, the remaining parameter estimate still had a significant substantive implication for the arrest rates.

Thus, as with the earlier samples, the results generated for this 1991-92 sample from the semiparametric mixed Poisson model also clearly support the "mixed" position. First, there was significant population heterogeneity in the propensity to commit criminal acts, and accounting for these differences was critical for estimating the genuine

Figure 8.3. Predicted Mean Arrest Charges, by Arrest Status at Prior Age: 1991-92 Sample



magnitude of the state dependence parameter. Second, even after controlling for the unobserved heterogeneity in the propensity to commit criminal acts, there was still a significant, positive, and substantively important relationship between the mean offending rate at a given age and whether or not an individual was arrested at the prior age. This importance of the state dependence effect is graphically depicted in Figure 8.3. Similar to the results for the 1981-82 and 1986-87 samples, the simple adjustment of moving from a 1-class model to the 2-class model (in the 1991-92 sample) had the largest effect on the reduction of the magnitude of the state dependence effect (a decrease of 0.178), whereas the change that resulted from moving from a 5-class model to a 6-class model was considerably smaller (a reduction of 0.023). Again, the fact that speed of decline in the magnitude of the state dependence parameter had significantly tapered-off was a substantive indication that the mixing distribution (i.e., distribution of unobserved heterogeneity) was adequately approximated with 6-points of support. Most of the 7- and 8-class local solutions for this data had the state dependence parameter right at 0.20 (i.e., nearly identical to the 6-class estimate).

Stage Two—The Parametric Random Effects Negative Binomial Model

Moving on to the second stage, here we present results from the parametric random effects negative binomial model with the 1991-92 sample and again compare the resulting estimate of this model to the estimate of the state dependence parameter presented above in the semiparametric mixed Poisson model. The results from the random effects model are presented in the first column of Table 8.12 (under Model 1). In

Table 8.12. Investigation of State Dependence Effects with Parametric Random Effects & NB1 Negative Binomial Models: 1991-92 Sample

	Model			
	1	2	3	4
Overall Intercept	-6.797	-7.622	-6.989	-32.806
	(55.29)	(58.35)	(53.46)	(19.50)
Latent Class Indicators				
LC1	---	0.539	0.539	25.906
		(11.71)	(11.71)	(15.27)
LC2	---	0.347	0.347	16.957
		(7.26)	(7.26)	(9.71)
LC3	---	1.103	1.103	24.856
		(23.58)	(23.58)	(14.62)
LC4	---	0.779	0.779	27.603
		(16.17)	(16.17)	(16.26)
LC6	---	1.054	1.054	27.744
		(20.43)	(20.43)	(16.18)
Age Effects				
Age	6.679	6.982	6.982	44.231
	(46.54)	(49.20)	(49.20)	(19.79)
Age Sq	-1.795	-1.873	-1.873	-14.603
	(45.79)	(48.41)	(48.41)	(19.84)
LC1 * Age	---	---	---	-37.304
				(16.60)
LC1 * Age-Squared	---	---	---	12.888
				(17.46)
LC2 * Age	---	---	---	-25.533
				(11.11)
LC2 * Age-Squared	---	---	---	9.248
				(12.31)
LC3 * Age	---	---	---	-34.981
				(15.54)
LC3 * Age-Squared	---	---	---	12.174
				(16.47)
LC4 * Age	---	---	---	-37.240
				(16.53)
LC4 * Age-Squared	---	---	---	12.443
				(16.79)
LC6 * Age	---	---	---	-39.522
				(17.49)
LC 6 * Age-Squared	---	---	---	13.627
				(18.40)
State Depen. Effects				
Art_{t-1}	0.533	0.417	0.417	0.188
	(25.08)	(20.02)	(20.02)	(9.33)
Ancillary Parameters				
ln(r)	4.093	15.937	---	---
ln(s)	4.761	16.570	---	---
ln(alpha)	---	---	0.634	0.461
Log-Likelihood	-32072.642	-31520.064	-31520.063	-30094.455

Note: Absolute values of t-statistics are in parentheses. Latent class indicator variables are from the results presented in Chapter 6.

this model, the boundary-value likelihood ratio test for the presence of significant unobserved heterogeneity (that tests the significance of the r and s ancillary parameters) once again indicated the presence of significant heterogeneity (χ^2_{01} =14.34; p-value = 0.000). Turning attention now to the estimate of the state dependence parameter in the parametric random effects model (with only a single set of age parameters), we find that he state dependence estimate in this model was 0.533. This estimate is nearly identical in magnitude to the estimate generated in the 1-class semiparametric Poisson model (which makes no correction for persistent individual differences). Once again we find that the parametric random effects model significantly over-estimates the state dependence parameter in comparison with the final estimate we arrived at in the 6-class semiparametric model (which was 0.206). Indeed the estimate from the parametric random effects model is more than twice the size of the estimate from the 6-class semiparametric model. Again, this is a serious discrepancy in the magnitude of the effect and highlights the critical need for the multi-method approach employed herein.

Stage Three—Incorporating Latent Class Indicator Variables

In an effort to further investigate the discrepancy between the estimates of the state dependence parameter in the Model 1 of Table 8.12 and the estimate from the 6-class model in Table 8.11, we again employ the use of the random effects and standard (NB1) negative binomial models. We build on the model specification used in Model 1 of Table 8.12 by including the set of binary latent class indicator variables. Model 2 in Table 8.12 is a parametric random effects model that includes the set of binary latent class indicators from the results presented in Chapter 7, and Model 3 is Table 8.12 is the

NB1 negative binomial model. To be consistent with the specification used in the two earlier samples, the adolescent-limited group, denoted as LC5 in the Chapter 7 analyses, was again used as the omitted reference group in these models.

Comparing the results of Models 2 and 3 in Table 8.12, we again find that the two models produce identical solutions. A boundary-value likelihood ratio test comparing these two models (χ^2_{01} value = 0.001; p-value = 1.000) indicates that including the set of five binary indicators removed the presence of significant unobserved heterogeneity (which was previously found to be significant in stage two). Again, we find that the models with the binary latent class indicator variables produce a smaller estimate of the state dependence effect than does the parametric random effects model presented in Model 1 of Table 8.12. The effect in Model 3, however, was still much larger (almost twice the size) than was the estimate we arrived at earlier with the 6-class finite mixture model (0.206). Thus, while the indicator variables do once again remove the presence of significant individual-level heterogeneity, the estimate of the state dependence parameter is still significantly larger than it should be (according to the 6-class finite mixture model).

The specification used in Model 4 incorporates a set of interaction variables between the latent class indicator variables and the age and age-squared variables. In this part of the analysis, we are testing whether the state dependence effect uncovered in stage two *over-estimates* the true state dependence effect because of the failure to adequately account for the heterogeneity in the effects of the age parameters across the latent classes. The reader will recall that previously in Chapter 7 it was determined that allowing the age parameters to vary over the latent classes resulted in a significant improvement in the

model fit. The NB1 model presented in Model 4 of Table 8.12 adds 10 parameters to the specification of the NB1 model found in Model 3.

The fourth numerical column of Table 8.12 contains the parameter estimates for this model. All of the latent class * age interaction variables in Models 4 were highly significant, which indicates that the offending trajectories of the 5 latent classes were significantly different from the adolescent-limited group. Note the massive increase in the log-likelihood value (+1426) that resulted from adding these ten parameters to the model.

For the purposes of this study, however, the importance of allowing for latent class-specific age parameters in the model is that it permits the successful recovery of a state dependence parameter estimate similar to that found in the semiparametric mixed Poisson model of Table 8.11. The state dependence parameter estimate in Model 4 of Table 8.12 was 0.188, whereas the estimate from the 6-class semiparametric model (of Table 8.11) was 0.206. In all three samples, the calculation of the estimate of the state dependence parameter was extremely sensitive to the accurate estimation of the "age effects" in the data. If the age effects are not adequately modeled, the state dependence covariate will absorb the effects, and the state dependence effect will appear to be significantly larger than it actually is.

Stage Four—Latent Class-Specific Models

The next stage of the analyses pertains to the models estimated on each latent class by themselves. Interest in these models concerns whether the overall estimate of the state dependence effect is generally reflective of the state dependence relationship

found within each latent class. In particular, we focus on the magnitude of the effect within the adolescent-limited group given the theoretical arguments of Moffitt (1993). Recall that in Chapter 7 an adolescent-limited group was discovered in the arrest data of the 1991-92 sample—this latent class was labeled LC5 in that chapter. The two classes with the highest arrest rates were the third and sixth latent classes. Each of these latent classes averaged around 27-28 arrest charges. The third latent class had an earlier and higher peak rate of arrest, whereas the sixth latent class did not reach its peak age of arrest until individuals were in their late twenties. The results from the NB1 negative binomial models estimated on each of the latent classes separately are found in Table 8.13.[22]

As indicated in Table 8.13, the estimates of the state dependence effects in all six of the latent classes were positive, but for the second latent class the estimate was not statistically significant at the .05 level (t-statistic = 1.14; p-value = 0.254). This latent class had 344 individuals (24%) in the sample assigned to it, and thus a significant, positive state dependence effect was estimated in the latent classes representing over 75% of the 1991-92 release sample. For the latent classes with significant positive estimates, the values ranged from 0.140 to 0.364. The two latent classes with the largest estimates were the fifth (0.314) and sixth (0.364) latent classes, but the 95% confidence intervals for the state dependence parameter estimates of these two latent classes completely overlapped with the 95% confidence intervals for the other three latent classes with

[22] In an identical finding to the 1981-82 sample, the boundary-value likelihood ratio tests of these NB1 models against the parametric random effect models (which allow for individual-level effects) resulted in a failure to reject the null hypothesis of no individual-level effects for all six latent classes. The value of χ^2_{01} in all six of the tests was equal to 0.00 (p-value=1.00), which indicates that the log-likelihood values for the NB1 model (which ignores individual differences) were all identical to the log-likelihood values of the random effects models.

Table 8.13. Investigation of State Dependence Effects with NB1 Negative Binomial Models: 1991-92 Sample, Latent Class-Specific Models

	Latent Class					
	LC1	LC2	LC3	LC4	LC5	LC6
Overall Intercept	-[illegible].097	-16.653	-[illegible]	[illegible].979	-31.815	-4.870
	(29.71)	(31.74)	(26.64)	(19.24)	(17.69)	(13.97)
Age Effects						
Age	7.096	19.678	9.033	6.794	42.914	4.532
	(27.37)	(31.85)	(27.56)	(19.90)	(17.94)	(12.77)
Age-Squared	-1.759	-5.636	-2.368	-2.101	-14.199	-0.949
	(26.17)	(32.01)	(27.04)	(20.45)	(18.15)	(11.00)
State Depen. Effects						
Arr_{t-1}	0.140	0.054	0.192	0.257	0.314	0.364
	(3.62)	(1.14)	(4.38)	(5.25)	(3.69)	(5.55)
Ancillary Parameters						
ln(alpha)	0.155	0.198	0.788	0.508	0.486	0.722
Log-Likelihood	-8193.382	-5400.409	-6520.795	-5151.611	-1772.791	-2984.292
N (Panel)	8469	6930	4598	4160	3068	2160
N (Observations)	396	344	224	211	159	100

Note: Absolute values of t-statistics are in parentheses.

significant estimates. Furthermore, a test of significant differences between the parameter estimates of each of the two latent classes with the largest effects and the parameter estimates of the other three groups failed to reject the null hypothesis of no differences. Similar to the previous obtained results for the 1981-82 and 1986-87 samples, after making allowances for sampling variation, the hypothesis that the five significant state dependence coefficients were all equal could not be rejected. Excluding the one latent class with a non-significant positive effect, the results in their entirety indicated that the overall effect seemed to accurately represent the state dependence relationship in the vast majority of the latent classes in the sample. The evidence here, as in the two previous samples, refutes the hypothesis that the state dependence effects in the adolescent-limited offender group are more pronounced (i.e., the evidence refutes hypothesis 8 of this study).

Stage Five—Post-Release Data Only

Finally, the last set of results presented in this chapter focus on the sensitivity analyses in which we limit the investigation by using only the values of the dependent variable from the post-release period. Here we examine whether any of the background characteristic variables are significantly related to the post-release arrest rate.

Results for four models (where only the post-release arrest data was used to construct the dependent variable) are presented in Table 8.14. The first model is a standard NB1 negative binomial model that completely ignores individual differences in the propensity to offend. The state dependence parameter estimate from this model is comparable to the 1-class parameter estimate presented in Table 8.11. Here we see that

Table 8.14. Investigation of State Dependence Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1991-92 Sample, Post-Release Data Only

	Model			
	1	2	3	4
Overall Intercept	-5.131	-6.205	-6.832	-13.711
	(7.14)	(8.53)	(9.59)	(6.52)
Latent Class Indicators				
LC1	---	---	0.258	1.972
			(5.44)	(0.78)
LC2	---	---	-0.373	-7.343
			(6.84)	(2.04)
LC3	---	---	0.829	-4.100
			(17.28)	(1.48)
LC4	---	---	-2.225	-64.418
			(15.68)	(2.72)
LC6	---	---	1.067	-3.745
			(19.71)	(1.29)
Age Effects				
Age	4.737	5.462	6.297	13.192
	(7.51)	(8.53)	(10.07)	(6.76)
Age Sq	-1.160	-1.330	-1.531	-3.212
	(8.50)	(9.58)	(11.31)	(7.17)
LC1 * Age	---	---	---	-3.041
				(1.33)
LC1 * Age-Squared	---	---	---	0.995
				(1.94)
LC2 * Age	---	---	---	8.592
				(2.55)
LC2 * Age-Squared	---	---	---	-2.454
				(3.15)
LC3 * Age	---	---	---	4.103
				(1.63)
LC3 * Age-Squared	---	---	---	-0.823
				(1.45)
LC4 * Age	---	---	---	71.255
				(2.86)
LC4 * Age-Squared	---	---	---	-20.001
				(3.06)
LC6 * Age	---	---	---	0.993
				(0.39)
LC 6 * Age-Squared	---	---	---	0.459
				(0.83)
State Depen. Effects				
Arr_{t-1}	0.546	0.357	0.219	0.056
	(18.08)	(10.51)	(7.28)	(1.86)
Ancillary Parameters				
ln(r)	---	1.790	---	---
ln(s)	---	1.996	---	---
ln(alpha)	0.672	---	0.411	0.301
Log-Likelihood	-14913.945	-14813.791	-14021.508	-13618.151
N (Panel)	13180	13180	13180	13180
N (Observations)	1434	1434	1434	1434

Note: Absolute values of t-statistics are in parentheses.

the estimate is of a similar magnitude to the estimate presented in Table 8.11 (0.546 versus 0.553).[23] The next model in Table 8.14 (denoted as Model 2) presents the estimates from a random effects negative binomial model applied to the post-release data. The model presented earlier (using the full available data) that is comparable to this particular model is Model 1 of Table 8.12. The state dependence parameter estimate in Model 2 here was 0.357, which is smaller than the estimate found in Model 1 of Table 8.12 (0.533). The substantive interpretation of this model, however, is identical to the earlier model based on the full data.

Next we consider Model 3 of Table 8.14, which is directly comparable to Model 3 of Table 8.12. The estimate of the state dependence parameter in this model was 0.219.[24] The estimate from the model making full use of all available data points was 0.417 (Model 3 of Table 8.12). Finally, the last model found in Table 8.14 is directly comparable to Model 4 of Table 8.12. Unlike the two earlier samples, using only the post-release data, in the 1991-92 sample (and properly accounting for the age effects) we still found a positive effect, but the estimate was now only marginally significant (estimate = 0.056; p-value=0.062). However, the effect was still positively related to the mean rate of offending. This is the first sensitivity analysis we have conducted in which the resulting substantive interpretation of the estimates was not identical to that obtained with the entire set of available data. It is important to keep in mind that only 8 years of post-release data were used here, and that many age-years of data (over 16000 data

[23] The parameter estimate directly comparable to the 1-class model would be calculated using a standard Poisson model. We re-estimated Model 1 of Table 7.9 using a standard Poisson model. The corresponding estimate for the state dependence variable in the Poisson model was 0.603 (t-statistic = 29.49).
[24] The model directly comparable to Model 2 of Table 8.12 (i.e., a random effects model with the latent class indicators) produced an identical solution to Model 3 of Table 8.14. Once the latent class indicators were included, the random effects version of the negative binomial model was not necessary

points) were excluded from the analysis here. The substantive importance of this finding is that researchers examining the evidence for state dependence versus population heterogeneity may arrive at different conclusions depending on the length of the time period over which their samples are studied.

We note that in all of the analyses in which we consider only the post-release data (in all three samples), the state dependence effects were smaller than were such effects calculated when using the full array of data for each sample. This leads to a possible interpretation that the state dependence effects may have been stronger in the years prior to incarceration in the CYA, and that the differences between the two periods may have been more pronounced for the 1991-92 sample. This is an interesting question (i.e., whether the state dependence effect is strongest before becoming a ward/parolee of the CYA). The answer to this question is beyond the scope of this chapter.[25] For now we leave this as a topic of possible future inquiry.

Finally, attention is now turned to the question of whether there are any significant predictors of the post-release arrest activity of the 1991-92 sample. The results of the models addressing this issue are presented in Table 8.15. As was the case with the two earlier (1981-82 and 1986-87) samples, the models with the 1991-92 data simply add the background variables to the model specifications used to estimate the models found in Table 8.14. Model 1 is the "naïve" NB1 model that ignores individual differences in the propensity to commit criminal acts. In this model, the following covariates were significantly related to the post-release arrest rate (they are in bold type

[25] This question would require generating entirely new data analytic files from based using the date of *first* admission to the CYA as dividing point, which is why it is not explored herein.

Table 8.15. Investigation of Subject Characteristic Effects With Parametric Random Effects & NB1 Negative Binomial Models: 1991-92 Sample, Post-Release Data Only

	Model			
	1	2	3	4
Overall Intercept	-5.447	-6.424	-7.057	-14.208
	(7.53)	(8.76)	(9.78)	(6.74)
Set of Control Variables	A	A	B	C
Background/CYA Vars.				
African American	0.152	0.200	0.029	0.046
	(3.28)	(3.56)	(0.64)	(1.02)
Hispanic	0.003	0.013	-0.010	0.021
	(0.06)	(0.22)	(0.22)	(0.47)
Other Ethnicity	-0.375	-0.389	-0.169	-0.142
	(4.36)	(3.94)	(2.04)	(1.73)
Family Violence	-0.051	-0.054	-0.008	0.008
	(1.26)	(1.12)	(0.21)	(0.20)
Par. Alc./Drug.	0.035	0.033	0.009	0.008
	(1.01)	(0.78)	(0.26)	(0.24)
Par. Crim.	0.046	0.027	0.031	0.015
	(1.31)	(0.64)	(0.91)	(0.44)
Sibling Crim.	0.043	0.050	0.019	0.010
	(1.43)	(1.38)	(0.67)	(0.36)
Neglect	0.125	0.138	0.033	0.045
	(3.56)	(3.29)	(0.96)	(1.33)
Control	0.059	0.053	0.059	0.087
	(1.18)	(0.91)	(1.20)	(1.77)
Abuse	-0.079	-0.079	-0.026	-0.029
	(1.94)	(1.60)	(0.66)	(0.73)
Sex Abuse	-0.069	-0.067	-0.043	-0.050
	(0.99)	(0.80)	(0.63)	(0.74)
Drug Abuse	0.076	0.081	0.041	0.024
	(2.18)	(1.95)	(1.22)	(0.71)
Gang Member/Ass.	0.028	0.016	0.040	0.028
	(0.72)	(0.34)	(1.04)	(0.74)
School Dropout	0.141	0.167	0.097	0.086
	(2.98)	(3.00)	(2.10)	(1.89)
Juvenile Court	0.011	-0.001	0.045	0.170
	(0.18)	(0.02)	(0.77)	(2.93)
First Commit	-0.187	-0.194	-0.027	0.007
	(5.78)	(4.93)	(0.85)	(0.21)
Comm. Off. Viol.	-0.081	-0.086	-0.076	-0.110
	(2.50)	(2.23)	(2.42)	(3.51)
Los Angles	-0.144	-0.096	0.002	-0.020
	(4.28)	(2.37)	(0.07)	(0.62)
Infraction Rate	0.263	0.375	0.187	0.180
	(3.37)	(3.36)	(2.25)	(2.20)
State Depen. Effects				
Arr_{t-1}	0.499	0.352	0.204	0.030
	(16.44)	(10.53)	(6.75)	(1.01)
Ancillary Parameters				
ln(r)	---	2.014	---	---
ln(s)	---	2.279	---	---
ln(alpha)	0.639	---	0.407	0.295
Log-Likelihood	-14804.210	-14750.155	-14000.990	-13590.728

Notes: Absolute value of t-statistics in parentheses. Sets of control variables are as follows:
Set A: Age, Age-Squared; Set B: Set A + Latent Class Indicators; Set C: Set A + Set B
+ entire set of interaction variables of the latent class indicators and the age variables.

in Table 8.10): African American ethnicity, Other ethnicity, neglected, drug abuse, school dropout, violent commitment offense, coming from Los Angeles County, and the DDMS infraction rate. The random effects negative binomial model essentially leaves the results from Model 1 unchanged. The random effects model assumes that the random effects are uncorrelated with the included covariates, an assumption that is highly tenuous. Thus, a better approach to testing whether any of these covariates are significantly related to the arrest rate (net of the effects of unobserved heterogeneity) is to use the latent class indicators to control for the unobserved heterogeneity. Models 3 and 4 include such indicators in the specification, and similar to the results obtained from the two previous samples, the inclusion of these indicator variables in the equations renders most of the significant covariates in Models 1 and 2 insignificant. In fact, in these two models (Models 3 and 4) there were only 3 covariates that were still significantly related to the post-release arrest rate: juvenile court commitment, violent commitment offense, and DDMS infraction rate. Clearly, determining what variables are significantly related to the arrest patterns in samples such as these requires employing the use of adequate controls for unobserved heterogeneity.

SUMMARY OF RESULTS

Having now completed the presentation of results for this chapter, here we briefly summarize the results obtained for each of the release samples. This chapter then concludes with a discussion of these results and how they provide evidence to support or refute the hypotheses promulgated in Chapter 3.

Summary of Results for the 1981-82 Sample

The analyses of the 1981-82 release sample began with a presentation of the semiparametric models estimated with varying numbers of points of support (from 1 point of support through 6 points of support). The results of the models presented in Table 8.1 indicated that accurately controlling for persistent individual differences was necessary for calculating non-spurious estimates of the state dependence parameter. The magnitude of the estimate of the state dependence parameter decreased by more than one-half between the model that specifies no controls for individual differences (1-class model) and the 6-class model. Yet, even after controlling for persistent individual differences through the use of six points of support, having been arrested during the prior age, *ceteris paribus*, significantly increased the frequency of arrest at the next age. In other words, the results favored the *mixed* position that allows for the important effects of both population heterogeneity and state dependence.

In Table 8.2, we examined two broad issues. First, we examined if we could reproduce the final estimate of the state dependence parameter calculated in the 6-class model of Table 8.1 with the parametric random effects negative binomial model. The parametric random effects model was found to produce a significantly larger estimate of

the state dependence parameter. Second, we examined if allowing the estimated age coefficients to vary over the latent classes allowed for an accurate recovery of the state dependence parameter estimate. After allowing the age coefficients to vary over the latent classes, we were able to almost identically reproduce the state dependence parameter estimate found in Table 8.1. The state dependence parameter is apparently very sensitive to shifts in the mean rate of offending among the latent classes, and the parametric random effects model in its basic specification does not directly estimate this heterogeneity. This type of heterogeneity, however, is explicitly accounted for the in the finite mixture approach of Nagin and Land (1993). Clearly, inaccurately accounting for the varying trajectories of offending over the age distribution will lead to an overestimation of the state dependence parameter. Similar to the findings of Bushway et al. (1999), we find that unaccounted for age effects are adept at masquerading as true state dependence effects.

Attention was next turned to an examination of whether the estimates of the state dependence effects varied across the latent classes. The results presented in Table 8.3, which were generated by estimating a NB1 model on each latent class separately, indicated that the state dependence estimates could be robustly estimated even within the latent classes. Importantly, the effect calculated within the adolescent-limited group was very similar to the effects calculated in all of the other groups, a finding that runs counter to the prediction of Moffitt (1993).

In the final stage of the analysis, we examined whether (1) any of the results would have changed if we only had access to the post-release data to construct the dependent variable and (2) if there were any covariates found to be significantly related

to the post-release arrest rates of this sample. For the most part, the substantive findings were reproduced using only the post-release data as the dependent variable. Furthermore, after controlling for persistent individual differences, we found that few variables were significantly related to the post-release arrest rate. Without controlling for the persistent individual differences, however, we would have come to the erroneous conclusion that there were many significant covariates (e.g., drug abuse, school dropout) related to the post-release arrest rate.

Summary of Results for the 1986-87 Sample

In stage one, the application of the semiparametric mixed Poisson model for the 1986-87 sample produced results consistent with the "mixed" position. Accounting for the significant population heterogeneity was critical in order to obtain the best estimate of the state dependence relationship. Failure to account for the unobserved heterogeneity would have resulted in a state dependence estimate that was significantly larger than the final estimate arrived at in the 6-class model presented in Table 8.6. Yet, even after controlling for the unobserved heterogeneity, there was still a significant relationship between the variable indicating arrest at a prior age and the arrest frequency at the current age. More specifically, the individuals arrested at the prior age had a significantly higher arrest frequency.

In the next stage of the analysis, we applied the random effects estimator and found the estimate of the state dependence relationship in this model to be much larger than the estimate found in the 6-class semiparametric model. However, after allowing the age parameters to vary over the latent classes, we were able to recover the same state

dependence estimate we found in the 6-class mixture model. We next tested whether the state dependence effect varied within the three latent classes. Except for the one latent class that was found to have a positive, yet non-significant effect, the state dependence estimates calculated within the latent classes were fairly similar. The estimate of the state dependence effect in the adolescent-limited group, however, was not found to be significantly larger (or smaller) than the other significant estimates. It was found to be similar to the other significant state dependence estimates once sampling variation was taken into account.

In the last section, we found the substantive results favoring the "mixed" position were unaltered by limiting the dependent variable to only the post-release years. Finally, the number of covariates that were significantly related to the post-release arrest rates of this sample was reduced to only a handful after controlling for the effects of unobserved heterogeneity though the use of the latent class indicators. Half of the significant covariates (e.g., drug abuse, school dropout) became insignificant after inclusion of the latent class indicators into the model.

Summary of Results for the 1991-92 Sample

The findings for the 1991-92 sample were consistent with the findings discussed above for the two earlier samples. In the first analysis stage, it was discovered that there was again a highly significant positive relationship between having been arrested at the prior age and the frequency of arrest at the current age. However, a significant portion of this relationship was subsequently found to have been the result of the population heterogeneity processes. Between a 1-class semiparametric mixed Poisson model (which

assumes homogeneity) and a 6-class model (which approximates the mixing distribution with 6 points of support), the magnitude of the state dependence parameter estimate decreased by 63%. Even so, after adequately controlling for unobserved heterogeneity, there was still a significant positive relationship between having been arrested at the prior age and the mean arrest rate at the current age.

In the next stage, we examined the estimate of the state dependence relationship using the parametric random effects negative binomial model. Consistent with the findings from the two earlier samples, the random effects negative binomial model produced a significantly larger estimate of the state dependence relationship. In fact, the estimate of the parametric random effects model was again more than twice the size of the estimate from the 6-class semiparametric random effects model. Yet, once again, adequately accounting for the diverse age effects across the latent classes allowed for the recovery of a nearly identical state dependence parameter estimate.

Models were next estimated within each of the latent classes. In five of the six latent classes, the estimate of the state dependence parameter was found to be significant and positive. For one of the latent classes, however, the positive coefficient failed to attain statistical significance. Examination of the differences in the state dependence estimates of the five latent classes with significant estimates failed to uncover any significant differences between them. The adolescent-limited group had a state dependence estimate that was similar (within sampling variation allowances) to the estimates of the other latent classes that had significant effects.

Finally, results presented in the fifth section indicated one finding consistent with the two earlier samples and one finding that differed from the two earlier samples.

Analyzing only the post-release data (and accurately accounting for the age effects) resulted in the finding that the state dependence parameter estimate was only marginally related to the post-release arrest rate. In the two earlier samples, the substantive findings were completely unaltered by the analysis of the post-release data exclusively. These sensitivity analyses were important because they highlighted the benefit of: (1) having access to data that has extensive coverage of the age distribution (rather than only having a few years worth of data); (2) the need to be careful about drawing conclusions based on data with limited time periods; and (3) the benefit of being able to replicate analyses on more than one sample. Similar to the two earlier samples, though, the analysis of the importance of the background characteristics in explaining the post-release arrest rates resulted in a finding that was compatible with the two earlier samples. That is, controlling for persistent individual differences, only a few of the covariates were found to be significant predictors of the post-release arrest rate. More than half of the significant predictors were reduced to non-significance once persistent individual differences were adequately controlled. This finding also highlights the importance of controlling for persistent unobserved heterogeneity in any analysis of data that attempts of evaluate the empirical adequacy of the three theoretical perspectives of concern in this study.

DISCUSSION

The purpose of this chapter was to carefully examine the positive association between past and subsequent offending. More specifically, the substantive focus of this chapter concerned the nature of the relationship between criminal activities at adjacent

ages while controlling for persistent individual differences. In Chapter 2 of this study, the etiological importance of this relationship was discussed. Three theoretical perspectives on the association were discussed, including the population heterogeneity explanation (represented by the theory of Gottfredson and Hirschi), the state dependence explanation (represented by the theory of Sampson and Laub), and the dual taxonomy explanation (represented by the theory of Moffitt).

Gottfredson and Hirschi (1990) argue that once differences in criminal propensity have been accurately controlled, the relationship between past and subsequent criminal activity should be reduced to non-significance (within sampling variation). According to Gottfredson and Hirschi (1990), the relationship between past and subsequent criminal activity is *spuriously* due to population heterogeneity in the propensity to engage in criminal activities.

Sampson and Laub (1993, 1997), on the other hand, argue that even after controlling for persistent individual differences in criminal propensity, there should still be a significant positive association between the levels of criminal activity at two points in time. Criminal activity at one point in time should (even after controlling the propensity to engage in criminal activity) still be positively related to subsequent criminal offending because such activities mortgage or cut off the future options of the offender and negatively alter the social bond (i.e., it negatively alters their local life circumstances). Sampson and Laub do not limit their theoretical argument to certain types of offenders (e.g., high- or low-rate), but rather they argue that their theory applies to all offenders, especially the population of serious offenders (such as those they used in developing and testing their theory).

In contrast to the first two theoretical perspectives, Moffitt (1993) argues for the importance of both population heterogeneity and state dependence process, but each is limited to only a single offender type. The behavior of the life-course-persistent offender type is argued to be theoretically governed by a static, population heterogeneity explanation, whereas the behavior of the adolescent-limited offender type is entirely governed by a state dependence explanation. According to this stream of theoretical insight, there should be a pronounced state dependence effect in the adolescent-limited group of offenders, and there should be a limited (or non-existent) state dependence effect in the life-course-persistent group of offenders. In other words, the dual taxonomy approach of Moffit envisions differential state dependence effects across the two offender types.

In Chapter 3, we reviewed the extant literature on this topic, and we concluded that further study of this topic was important and warranted because of two key limitations of the previous literature. First, the importance of the state dependence perspective within the population of high-risk offenders has been questioned, but empirical investigations of samples of such offenders have been scarce. Second, the validity of the findings regarding the state dependence effect in the extant literature have been questioned on methodological grounds because the vast majority of studies concerned with this issue have relied entirely on the parametric random effects model to analyze data. This model assumes that both the mixing distribution (of the unobserved heterogeneity) follows a specific parametric distribution and that the offending process has been observed prior to initiation (the initial conditions assumption). With the exception of the study by Bushway et al. (1999), the majority of the previous studies have

not examined or tested whether obtained results were robust with respect to the method of analysis. The results presented in this study are the first application of the multi-method approach of Bushway et al. (1999) to data on the serious offender (or very high-risk) population.

In direct response to the calls for further investigations of this key theoretical issue by Bushway et al. (1999), Brame et al. (1999), and Nagin and Paternoster (2000), this study set out to examine four hypotheses concerning the relationship between past and subsequent criminal offending behavior using data collected on three samples of serious youthful offenders. The first hypothesis examined is this study was:

> **H_5**: There will be a statistically significant positive association between past and subsequent offending behavior.

The results presented in this chapter clearly support this hypothesis. Ignoring individual differences in the propensity to offend, the relationship between criminal offending at adjacent ages was found to be positive and highly significant. The estimate of this relationship was 0.857 (t-statistic = 86.28) in the 1981-82 sample, 0.619 (t-statistic = 51.58) in the 1986-87 sample, and 0.553 (t-statistic = 40.25) in the 1991-92 sample. The results concerning this hypothesis were important for establishing a baseline estimate of the relationship between past and subsequent offending in these three samples.

Again, the dispute regarding the results with respect to the first hypothesis centers not on the existence of the relationship (all parties to this dispute agree to the existence of a significant positive association), but rather on the interpretation of this relationship

The next three hypotheses that guided the research in this chapter centered on the dispute among the three theoretical perspectives over this issue:

> H_6: After controlling for persistent individual differences in criminal propensity, the association between past and subsequent offending will be reduced to a nonsignificant level (Gottfredson and Hirschi).
>
> H_7: After controlling for persistent individual differences in criminal propensity, the association between past and subsequent offending behavior will be reduced in magnitude but will still be positive and statistically significant (Sampson and Laub).
>
> H_8: The association between past and subsequent offending behavior will be nonsignificant for the life-course-persistent (high criminal propensity) group(s), while the effect should be substantial and significant for the "adolescent-limited" (or adolescent peaked) group (Moffitt).

The results presented in this chapter overwhelmingly support the seventh hypothesis (H_7), and largely fail to support the hypotheses delineated as H_6 and H_8. After accounting for the population heterogeneity in the propensity to engage in criminal activities nonparametrically in the semiparametric mixed Poisson model of Nagin and Land (1993), the estimates of the state dependence relationship in this chapter were 0.404 (t-statistic = 38.82), 0.245 (t-statistic = 19.18), and 0.206 (t-statistic = 14.59) in the 1981-

82, 1986-87, and 1991-92 samples, respectively. The estimates from the parametric random effects negative binomial model (after properly accounting for the age effects in the data) were 0.402 (t-statistic = 0.402), 0.248 (t-statistic = 13.90), and 0.188 (t-statistic = 9.33) in the 1981-82, 1986-87, and 1991-92 samples, respectively. Thus, after accounting for persistent unobserved heterogeneity in the propensity to engage in criminal activities (as measured by arrest data) through both parametric and nonparametric methods, there was still a significant positive relationship between having been arrested at the prior age and the frequency of arrest at the current age. Results such as these clearly support the seventh hypothesis, and explicitly refute hypothesis six. That is, even after accounting for population heterogeneity in criminal propensity within the three samples, there was still a significant positive relationship between criminal activity patterns at adjacent ages. Statistical tests for unaccounted individual differences failed to reject the null hypothesis of no individual-specific effects, and thus the remaining state dependence effects uncovered in the models cannot be simply argued to be spuriously due to persistent unobserved differences.

It should be noted, however, that it was absolutely critical to adequately control for the differences in criminal propensity when estimating the relationship between past and subsequent criminal activity. There was a very large decrease in the magnitude of the state dependence relationship after controlling for persistent individual differences. In fact, there was a consistent 50-60% reduction of the magnitude of the state dependence parameter between the initial baseline estimate (from the 1-class semiparametric mixed Poisson model) that makes no allowances for individual differences in the propensity to commit criminal acts and the final estimates arrived at after accounting for the

unobserved heterogeneity (and age effects). Clearly, accounting for individual differences is critically important when examining the relationship between past and subsequent criminal activity, even within these three samples of serious offenders. Even so, it should be kept in mind that there still remained a significant relationship between criminal offending at adjacent ages in the final models presented in this chapter. Graphs of the predicted arrest rates for those who had and had not been arrested at the prior age were used to display the final estimates. They implied that there was a substantively meaningful and important relationship for these covariates, even if they were significantly reduced from those obtained from initial estimates.

With respect to the eighth hypothesis examined in this study, the results of the models estimated within each latent class failed to uncover significant differential state dependence effects that were stronger in the adolescent-limited group. There were two latent classes found to have positive state dependence effects that failed to attain statistical significance (at conventional levels). For the other 16 latent classes with significant positive effects, however, the class-specific state dependence estimates were not found to be significantly different from one another after taking possible chance sampling variation into account. The significant positive effects in the three adolescent-limited groups were found not to be of a significantly greater magnitude when compared with the significant estimates in the other 13 latent classes. Thus, the evidence examined in this study failed to support the eighth hypothesis. The latent classes that most closely resembled (in a relative sense at least) the life course persistent group (i.e., their criminal offending extended further into adulthood) also had significant positive estimates that were very close to the estimates found within the adolescent-limited groups. No evidence

was found in any of the three samples indicating that the state dependence effects were more pronounced in the adolescent-limited group. The state dependence effect uncovered in the data appeared to be a *general* effect that applied to all offender types, and not to just a specific offender type such as the adolescent-limited group. The fact that similar positive (significant) state dependence effects were found in sixteen of the 18 latent classes sheds a considerable degree of empirical doubt on the dual taxonomy perspective of Moffitt (1993). This evidence appears to lend further empirical support to the major tenets of the (mixed) theory of Sampson and Laub (1993) who argue that the state dependence effect should be robust across offender types (i.e., apply to all offenders), rather than any specific type of offender (as claimed in Moffit's theory).

The results presented in this chapter also resonate with a significant methodological theme on this topic. Previous research has discussed how reliable conclusions regarding the importance of state dependence processes (versus population heterogeneity) are contingent on the proper specification of the underlying mixing distribution (distribution of unobserved heterogeneity). This is the benefit of the multi-method approach—to the degree one can replicate a finding across various methods that make different and/or more/less stringent assumptions, the tenability of the finding(s) become more reliable. The use of a single method of analysis leaves the results of a study in a "gray" area, where conclusions often will undoubtedly (and justly) be viewed with a healthy degree of skepticism (Bushway et al. 1999; Nagin and Paternoster 2000). Results shown to be robust across different methods and model specifications, however, will be given greater credence.

However, the results obtained in this study clearly speak to the need to not only apply the multi-method approach, but also of the need to adequately model the "age effects" in these data. The results for all three samples obtained in Chapter 7 clearly showed both an overwhelming change in the nature of criminal offending patterns across the age distribution as well as a considerable degree of heterogeneity in the offending patterns among the latent classes. The findings obtained here were not indicative of a common trajectory that was merely differentiated in terms of the mean arrest rate. There was considerable heterogeneity in the nature of the estimated age parameters among the latent classes. The results of the analyses presented in this chapter clearly indicate that a failure to accurately capture the age effects within a sample of data will lead to a severe overestimation of the estimated state dependence effect. According to the results presented herein, this point is of fundamental importance from a methodological standpoint because it highlights the need for researchers to think critically about the proper specification necessary to accurately model the age effects. For example, we re-estimated the semiparametric mixed Poisson model, and constrained all of the age parameters to be equal across the latent classes such that

$$\ln\left(\lambda_{it}^{j}\right) = \left(\beta_0 + \bar{\varepsilon}_j\right) + \left((age_{it}/10) * \beta_{age}\right) + \left(\left(\left(age_{it}^2\right)/100\right) * \beta_{age^2}\right) + \left(arr_{t-1} * \beta_{arr_{t-1}}\right). \qquad (4)$$

The estimates of the state dependence relationship from these models were nearly identical to the those estimated in the parametric random effects models where the age affects were only accounted for through the use of two overall terms (i.e., Model 1 of

Tables 8.2, 8.7, and 8.12). The estimates from these semiparametric models were 0.60 (1981-82), 0.43 (1986-87), and 0.42 (1991-92) respectively, which are almost identical to the random effects specification. These coefficients still represent an over-estimation of the state dependence effects compared to that obtained when the age effects were properly modeled. Thus, while agreeing that the multi-method approach is a very important technique that should be adhered to when addressing this topic of research, simply reproducing the results across the different methods still does not mean the estimates are correct. We could accurately reproduce the nearly identical "over-estimates" of the state dependence parameter in both the parametric and semiparametric approaches to the mixing distribution. The methodological contribution that the results presented in this study suggest is that it is critical to ensure that the age parameters are adequately modeled in the data because unaccounted for variation in the age parameters appears to be quite adept at masking genuine state dependence effects (Bushway et al. 1999). Had we not accounted for the diverse age effects across the latent classes, the estimates of the state dependence parameter would have *doubled* in magnitude (in both the parametric and semiparametric mixed Poisson models). Without the flexibility of the semiparametric mixed Poisson model of Nagin and Land (1993), this realization would never have come about. The only way to account for the temporal shifts in the mean arrest rates was to either: (1) implicitly allow the age effects to vary across the latent classes as in the semiparametric mixed Poisson model specified in equation (1); or (2) to have access to the latent class indicators which could be interacted with the overall age parameters. Thus, although Bushway et al. (1999) note that, "the addition of time controls to random effects models is a very simple task," accounting for the class-specific

effects cannot be accomplished in the parametric random effect models by themselves. Therefore, results obtained from these models should be viewed with more than a healthy degree of skepticism. The direct benefit of the semiparametric mixed Poisson model is that its flexibility allows for building the variation in the age parameters directly into the specification. Perhaps this is only relevant for the high-rate groups studied here, but more than likely, it will be valid any time there are significant differences in the age parameters and the estimated age parameters do not accurately reflect the actual age parameters in the high rate groups. The results obtained here indicate, quite emphatically, that the state dependence variable will absorb the effect of unaccounted for temporal, age-based variation that is unaccounted for in the model.

For the most part, however, the evidence presented in this chapter overwhelmingly favors the *mixed* position that allows for the general importance of both population heterogeneity and state dependence processes. The association between past and subsequent criminal activity cannot be simply argued to be a spurious artifact of population heterogeneity in criminal propensity. Yet at the same time, a large portion of the obtained relationship did appear to be a consequence of heterogeneity in the propensity to engage in criminal acts (i.e., the evidence does not favor either the pure population heterogeneity nor pure state dependence positions). Thus, in general, the results obtained in this chapter simply do not resonate with the strong static explanations embodied in the static population heterogeneity theories such as those offered by Gottfredson and Hirschi (1990) or Wilson and Herrnstein (1985).

To bring these results back into perspective with the overarching theme under which this study has been framed—the dual processes of continuity and discontinuity

(change) of criminal offending patterns across the life course—the results in this chapter clearly lend empirical support to the idea that both continuity and change are necessary for explaining the etiology of criminal offending across the life course; this appears to be true even within the serious offender population (and with a dependent variable based on arrest data). In the final chapter to follow, we conclude by considering the general overarching theme of this study and the implications of the results presented in this study for that theme. The final chapter concludes with a discussion of the limitations of this study and possible directions for future research on the issues examined in this study.

CHAPTER 9

CONCLUDING REMARKS

INTRODUCTION

This study began by discussing three questions critically important to the study of crime. We noted how two of the questions could be described with respect to the age-crime curves among latent classes of offenders and that the other question could be described with respect to the relationship between criminal activity at two adjacent ages. This naturally led to the formulation of two substantive chapters (7 & 8). Given that the results of each substantive chapter have been fully summarized, reviewed, and discussed already, this final chapter will contain concluding remarks that focus on the general theoretical and policy implications of the findings of this study and possible directions for future research.

THEORETICAL IMPLICATIONS

In a nutshell, the old adages, "you can't unscramble eggs" and "a leopard never changes its spots," describe the fundamental over-arching issue addressed in this study—the relevance of change in criminal behavior over the life course. Three broad theoretical frameworks were examined in this study. Each framework offers different predictions with respect to the possibility of behavioral change in the life courses of criminal offenders. As discussed in Chapter 2, the theoretical controversy between these three theoretical frameworks largely boils down to a single question: how stable or inflexible are individual differences in the propensity to engage in criminal/antisocial activities

across the life course? Or stated differently, is *change* possible in the lives of serious criminal offenders? Because each theoretical perspective envisions the stability of criminal propensity very differently, each makes different predictions regarding (1) the relationship between age and crime and (2) the relationship between past and subsequent criminal activities.

This study examined the above two different sources of "change." First, the issue of whether the relationship between age and crime was invariant across the latent classes (i.e., stable) or whether there were variable between-class differences over time (i.e., change over time) was examined in Chapter 7. Second, the issue of whether past criminal activity is related to subsequent criminal activity after controlling for persistent individual differences was examined in Chapter 8. The findings presented in Chapters 7 and 8 provided resounding support for the notion that behavioral change is extremely important to the explanation of the criminal offending and arrest patterns of serious youthful offenders. With respect to the first source of change, between-group differences were shown to be highly variable over time—between-group differences were stable only through early adolescence, and then during adulthood such differences were largely *instable* and variable. Also, even after accounting for persistent unobserved heterogeneity in the propensity to engage in criminal activities (through both parametric and nonparametric methods), there was still a significant positive relationship between having been arrested at the prior age and the frequency of arrest at the current age. In other words, having been arrested at the prior age appears to have changed the frequency of offending at the subsequent age. The broad substantive implications of these results are that behavioral *change matters* even in the lives of serious offenders, and even in the

lives of the most persistent serious youthful offenders too (who arguably would be the most prone to stability in the spirit of Moffitt's "life-course-persistent offender"). It is important to remember that these substantive findings were documented across three separate samples, which poses a considerable problem for any rival hypothesis suggesting that this pattern represents a statistical anomaly or fluke. Thus, the observed findings lend considerable credence to notion that behavioral change is a critically important factor for the etiological explanation of the observed patterns of criminal arrests across the life course.

There are three main substantive conclusions to be drawn from this study that are related to both the over-arching theme of continuity/discontinuity in criminal arrest patterns and also important for any etiological explanation of criminal offending patterns across the life course. First, the findings observed in this study speak directly to what Sampson and Laub (1992) referred to as the overstatement and/or misinterpretation of antisocial continuities across time (see also Loeber and Stouhamer-Loeber 1998). The findings presented in Chapter 7 clearly indicate that the continuity of arrest patterns will be much stronger when the measurement periods are closer in time. That is, there is much more stability (including between-group stability) displayed within shorter periods of time (e.g., 2 years) compared to longer periods of time. Indeed, the vast majority of the available datasets in criminology typically measure incidents of crime and arrest over very short periods of time. As noted by Cohen and Vila (1996: 147), "consistency of behavior might depend on the time scale selected for analysis." The empirical results of this study indicate a considerable amount of support for this contention. Continuity appears to be much stronger when the measurement points are closer together in time.

Discontinuity (change), on the other hand, is much more apparent when long-term criminal offending or arrest patterns are examined. Further, the results presented herein also indicate that some segments of the age span may indicate more continuity in behavioral patterns (e.g., early to mid-adolescence) than other segments of the age span (e.g., adulthood). Thus, while it is important for theory to recognize that there is continuity in behavioral patterns over time, the possible temporal nature of this continuity should be noted.

Second, the findings observed herein indicate that there is a significant amount of heterogeneity in the longitudinal arrest patterns of serious youthful offenders. Examination of both the observed average total arrest charges and the observed and predicted arrest trajectories of each latent class indicated that there was a significant amount of heterogeneity even in this select extreme segment of the offender population. Thus, these results appear to bolster the contentions of Sampson and Laub (1992, 1993; Laub and Sampson 2001) that there is far more heterogeneity in longitudinal criminal arrest and offending patterns than previously thought. It is important to specifically highlight that the type of heterogeneity to which we speak of here is heterogeneity in the patterns of criminal arrest over extended periods of the life course. As clearly shown in the findings presented in the graphical depiction of the arrest trajectories displayed in Chapter 7, there are periods of time when the distinctions between the arrest trajectories of the various latent classes were blurred or even non-existent. Over more extended periods of time, however, the trajectories were vastly distinct.

Third (and related to the second point noted above), during late adolescence and adulthood there appears to be large variations in criminal arrest patterns that can not be

argued as simply the long-term consequences of childhood propensities (Sampson and Laub 1992). As noted by LeBlanc and Loeber (1998: 131), "against the backdrop of [relative] continuity, studies also show large within-individual changes in offending, a point understressed by Gottfredson and Hirschi (1987)." The findings presented in Chapter 7 indicate that there were varying rates of developmental change in the arrest trajectories of the latent classes. For the most part, the latent classes appear to have discontinued their antisocial activities at very different paces and ages (see also Cernkovich and Giordano 2001). Among the identified latent classes, the "desistence process" did not begin at the same age, nor did it take place at the same relative pace. The results presented in this study thus add further empirical support to bolster the contentions of Sampson and Laub that "intra-individual change is widespread even among a large group of individuals labeled as serious, persistent youthful offenders and possessing all of the risk characteristics that many believe are enduring and stable across the life course" (Laub and Sampson 2001: 58).

In short, the main theoretical implications of this study suggest that the processes of both continuity and change (discontinuity) are important in any etiological explanation of criminal offending patterns across the life course (see also, Paternoster et al. 1997). Ignoring either of theses processes, or viewing them as polar opposites on a continuum will lead to inadequate explanations of criminal behavior (see Horney et al. 1995). Both processes are clearly relevant, and indeed the critical focus for future empirical research is determining the precise causal nexus behind each of these processes. Unfortunately, it is at this point in this study where the limits of the data employed here preclude its usefulness for examining such issues.

SUGGESTIONS FOR FUTURE RESEARCH

The reader will recall that in Chapter 2 we noted that this study was not able to test the specific causal structures of a particular theory or set of theories, but rather we stated that our study would present an empirical evaluation of the precise longitudinal implications of three leading criminological theoretical perspectives. In the end, the examination of the issues addressed in this study has led to the conclusion that the evidence obtained here lends considerable support to the implications of the theoretical perspective of Sampson and Laub (1993), and largely refutes the direct empirical implications of the perspectives of both Gottfredson and Hirschi (1990) and Moffitt (1993). The findings presented in this study, however, have only shown that behavioral change is evident among serious offenders and that the weight of the evidence favors the theoretical implications of the work provided by Sampson and Laub (1993). These findings, however, can by no means be construed as evidence for support of the causal structure of Sampson and Laub's criminogenic theory as outlined in their book. Again, the results of this study were merely consistent with the longitudinal implications of their theory—that is, there was heterogeneity in the criminal propensity between individuals. We found also that there was considerable post-adolescent heterogeneity in the arrest rates of offenders that cannot be explained as purely a consequence of earlier individual differences. Similarly, we found that there was a significant state dependence relationship between criminal offending at a prior age and the level of offending at the current age (even after controlling for individual differences in criminal propensity). Indeed, the limitations of the data employed in this study for assessing the causes of crime necessitate raising the issue of possible avenues for future research.

The data employed in this study, and the findings observed herein, cannot be used to answer the four critical questions raised by Nagin and Paternoster (2000) believed to be of critical importance for understanding the population heterogeneity-state dependence debate. First, what are the specific causal mechanisms underlying the individual differences in the propensity to commit criminal acts? Second, what are the specific positive and negative salient life events that lead individuals both into and out of the criminal lifestyle? Third, what are the specific causal processes underlying the desistence process? Finally, what processes determine both the availability of prosocial opportunities and whether or not an individual will take advantage of these opportunities?

A critical need for the discipline of criminology awaiting future research is to determine both theoretically and empirically the precise etiological mechanisms that are the driving force(s) behind the changes displayed in the nature of offending trajectories (see also Bushway et al. 2001; Laub and Sampson 2001). Particularly critical in importance is the task of determining why offenders, who have shown a pronounced proclivity (albeit a varying one) to engage in criminal activities for a significant segment of their life span would suddenly begin to decrease their offending in adulthood? Except for the path breaking work by Hirschi (1969) and Sampson and Laub (1993) on social control theory, there is relatively little theoretical or empirical research bearing directly on this issue. As is evident here in the lives of even the most serious offenders in the population, behavioral change occurs, and it occurs earlier for some individuals/groups than for others. The theoretical and public policy implications of the need to identify the sources of prosocial behavioral change among the serious offender population cannot be overstated.

Finally, even though it is clear that the chronic offenders within our samples appear to have been on a path of "desistence," exactly what kind of lives they actually lead in their thirties, forties, and thereafter is largely unknown at this point (Laub and Sampson 2001). Do the majority of these individuals lead highly marginalized lives full of alcohol and drug abuse problems, unemployment, and marital discord? Unfortunately, questions such as these cannot be answered with the data we have utilized here. Nonetheless, such questions remain ripe for consideration in future research. Having discussed the theoretical implications of our research and some possible directions for future research, we now conclude this study with a discussion of the possible policy implications of our analyses.

IMPLICATIONS FOR PUBLIC POLICY

Given the extremely high failure rates reported in Chapter 6, the pessimist who reads this study will argue that the benefits of institutional placement in the CYA appear to be very discouraging, at least in the short term. Upon reflection, we should not have expected low or moderate recidivism rates among our three samples of active offenders. Given the fact that the CYA is stocked with dedicated employees who by and large work very hard to rehabilitate and support wards under their supervision, and the program provides a variety of educational, treatment, training and supervisory services (outlined in Chapter 4), why shouldn't we expect low failure rates? First, we would do well to remember that these wards represent the worst 5% of the youthful offender population in the state. The case history records of these active offenders are considerably worse than any we have previously seen. For this and other groups of active offenders, we find no

consistent evidence to support the commonly held expectation that policy changes which increase the probability of arrest, severe punishment and the average length of sentence will significantly deter the likelihood of subsequent criminal behavior. The "get tough on crime" movement began in the mid-1970s as a justification for establishing an increase in severe sentencing decisions because of the failure of the rehabilitation programs (in vogue during the 1960s) that were supposed to reduce recidivism rates. The changes in CYA policy in the 1980s and early 1990s (outlined in Chapter 4) that have had the effect of increasing the average length of institutional stay over time do not appear to have improved the post release behavior of parolees as documented in Chapter 6, at least in the short-run. The simplest explanation for this finding is that the menu of education, treatment, training and supervisory services in place could not overcome or appreciably reduce the powerful forces influencing offenders to continue their criminal offending behavior.

The optimist who reads this report will focus on the long-term relationship observed between age and crime. While we have found little evidence to support Gottfredson and Hirschi's notion that the relationship between age and crime is invariant across latent classes over time, the idea that criminal behavior does not decline with age among active offenders advanced by their critics (e.g. Blumstein, Cohen, and Farrington 1988a, 1988b) has no support here. Recall that Blumstein and his colleagues contend that the rate of offending among active offenders reaches a peak level and then assumes a relatively constant rate. Over, the long term, here we observed that the arrest trajectories for every latent class derived from our active offender samples decline with age. The timing of the desistence patterns in our samples suggests to us that this is most likely due

to the processes of developing and strengthening social bonds identified in the work of Sampson and Laub or perhaps due to maturation rather than to a lagged beneficial institutional treatment effect. Control theories recognize important changes that naturally occur over the life course that reduce the likelihood of committing criminal acts. According to Gottfredson and Hirschi (1990:256), policies that do not consider these highly predictable circumstances are likely to mistake natural changes for program effectiveness and to waste considerable resources "treating" people without benefit to themselves or society. The decline in crime with age over time in each of the latent classes suggests that the maximum effect of selective incapacitation as a means of reducing crime in society should be focused on the age just prior to the rapid onset and peaking of criminal offending. We note, however, that it usually takes awhile to accumulate a criminal record sufficient to justify imprisonment. Often individuals are beyond the peak age of crime once they accumulate such a record.

The serious offenses that are of the greatest concern to society (e.g., index crimes) were, in fact, most frequently committed by members of our samples when they were relatively young (e.g., ages 14-22). That is, the data presented in this study showed a marked relationship between age and crime even for the serious criminal offenses that are the intended targets of selection incapacitation policies such as "Three Strikes." In fact, over time, the individuals in this sample became increasingly more likely to have been arrested for a drug-related offense than any other type of offense. For example, the members of the 1981-82 sample accrued 166 robbery charges in 1983. In 1999, they accrued only 18 arrest charges for robbery offenses. For drug-related offenses, on the other hand, they accrued 396 charges in 1983 and still accrued 299 charges for such

offenses in 1999. Because of their long histories of past involvement in serious crime, however, these individuals become prime candidates for selective incapacitation as "third strike" offenders at the approximate point in the age-crime curve where they no longer pose as grave a danger to the long-term public safety. The trends in our data suggest that the decline in crime with age is going to continue in the coming years, which (extrapolating to the general population of criminal offenders) indicates that there are going to be a fair number of offenders in prison who pose relatively little risk to society. If many of these individuals no longer represent a serious danger to society because they have "aged out" of or are in the final stages of aging out of serious crime, then there are potentially enormous social and economic costs to be paid for incarcerating them at later ages. Further, in order to finance the massive increases in state prison populations, legislatures have been forced to divert money from discretionary line items in state budgets—education, welfare, medical care, mental health services and child care. The fear among critics of "get tough on crime" policies is that money is being diverted away from the very same institutions that have traditionally played a crucial role in either preventing some individuals from engaging in serious crime in the first place or in helping individuals to desist from the criminal lifestyle. Ironically, the worry here is that higher incarceration rates may serve to set in motion a spiraling effect that in the long-term could push crime rates to rise rather than to decline.

APPENDIX A

OFFENSES SERIOUSNESS HIERARCHY

& CATEGORIZATION OF OFFENSE TYPES

Table A.1. Table of Offense Seriousness Looping Order and Categorization of Offense Types

Offense Type	Seriousness Loop Order	Serious Offense	Overall Category	Specific Category
Murder, First Degree	1	Yes	Serious Violent	Homicide
Murder, Second Degree	2	Yes	Serious Violent	Homicide
Manslaughter	3	Yes	Serious Violent	Homicide
Vehicular Manslaughter	4	Yes	Serious Violent	Homicide
Forcible Rape	5	Yes	Serious Violent	Forcible Rape
Attempted Murder	6	Yes	Serious Violent	Aggravated Assault
Aggravated Assault	7	Yes	Serious Violent	Aggravated Assault
Robbery (Enhanced--e.g., armed)	8	Yes	Serious Violent	Robbery
Robbery (Unspecified)	9	Yes	Serious Violent	Robbery
Robbery (Unenhanced)	10	Yes	Serious Violent	Robbery
Robbery Public Conveyance	11	Yes	Serious Violent	Robbery
Attempted Robbery	12	Yes	Serious Violent	Robbery
Extortion/Kidnapping	13	Yes	Serious Violent	Kidnap/Extortion
Child Molestation	14	Yes	Serious Violent	Child Molestation
Sodomoy/Forced Oral Copulation	15	Yes	Serious Violent	Sodomy/Oral Cop.
Discharge Weapons	16	Yes	Serious Violent	Weapons Discharge
Assault & Battery	17	No	Violent	Simple Assault
Miscellaneous Assault	18	No	Violent	Simple Assault
Burglary, First degree	19	Yes	Serious Property	Burglary
Burglary (Unspecified)	20	Yes	Serious Property	Burglary
Burglary, Second degree	21	Yes	Serious Property	Burglary
Attempted Burglary	22	Yes	Serious Property	Burglary
Auto Theft	23	Yes	Serious Property	Auto Theft
Grand Theft	24	Yes	Serious Property	Theft Major
Receiving Stolen Property	25	Yes	Serious Property	Theft Major
Forgery/Checks	26	Yes	Serious Property	Theft Major
Arson	27	Yes	Serious Property	Arson
Sales--Narcotics	28	Yes	Serious Drug	Drug Sales
Sales--Dangerous Drugs	29	Yes	Serious Drug	Drug Sales
Sales--Marijuana	30	Yes	Serious Drug	Drug Sales
Possession--Narcotics	31	Yes	Serious Drug	Drug Poss /Poss. For Sale
Possession--Dangerous Drugs	32	Yes	Serious Drug	Drug Poss./Poss. For Sale
Possession--Marijuana	33	Yes	Serious Drug	Drug Poss./Poss. For Sale
Weapons Possession	34	No	Fel. Weapons Poss.	Felony Weapons Poss.

Table A.1 (Continued).

Offense Type	Seriousness Loop Order	Serious Offense	Overall Category	Specific Category
Possession of Destructive Devices	35	No	Violent	Simple Assault
Disturbing the Peace/Carrying a Concealed Weapon	36	No	Other Residual	Misc. Misd.
Miscellaneous Felony	37	No	Other Residual	Felony Other
Accessory	38	No	Other Residual	Felony Other
Escape From a Secure Facility	39	No	Other Residual	Escape
Escape from Juvenile Secure Facility	40	No	Other Residual	Escape
Petty Theft	41	No	Property	Theft Minor
Reckless Driving with Injury	42	No	Other Residual	Felony Other
DUI with Injury	43	No	Other Drug	Other Drug
Statutory Rape	44	No	Other Residual	Other Sex Offense
Pimping/Prostitution	45	No	Other Residual	Other Sex Offense
Obscenity	46	No	Other Residual	Other Sex Offense
Contributing to the Delinq. of a Minor	47	No	Other Residual	Other Sex Offense
Misc. Sex Offense	48	No	Other Residual	Other Sex Offense
Malicious Mischief/Vandalism	49	No	Other Residual	Misc. Misd.
Trespassing/Loitering	50	No	Other Residual	Misc. Misd.
Auto Tampering	51	No	Other Residual	Misc. Misd.
Misc. Drugs Offenses	52	No	Other Drug	Other Drug
Under the Influence of Controlled Substance/Drugs	53	No	Other Drug	Other Drug
Sniffing ("Huffing" Paint)	54	No	Other Drug	Other Drug
Drunk/Disorderly Conduct	55	No	Other Drug	Other Drug
Miscellaneous Misdemeanor	56	No	Other Residual	Misc. Misd.
Traffic Offenses (eg. Driving w/o License)	57	No	Traffic/Vehicle	Traffic
Parole Technicality--AWOL	58	No	Parole Tech	Parole Tech
Parole Technicality--Other	59	No	Parole Tech	Parole Tech
Parole Technicality--Drugs	60	No	Parole Tech	Parole Tech
Parole Technicality--Gangs	61	No	Parole Tech	Parole Tech
Miscellaneous Minor Offenses	62	No	Other Residual	Misc. Misd.
Probation Violations	63	No	Probation Viol	Probation Viol.
Placement Failure	64	No	W&I Status	W&I
Incorrigible	65	No	W&I Status	W&I

APPENDIX B

OBTAINING MORTALITY DATA FROM THE CALIFORNIA DEATH STATISTICAL MASTER FILES

Mortality data on the subjects in the three CYA release samples were extracted from the Death Statistical Master Files (DSMF) of the California Department of Health and Human Services (DHS). The DSMF files are based on the death certificates completed by either the presiding physician at the time of death, or in the case of sudden or unexpected deaths such as homicide, suicide, or drug overdose, the coroner or medical examiner investigating the deaths. There is one DSMF file for each year. For example, all of the deaths that occurred between January 1, 1990 and December 31, 1990 would be included in the 1990 DSMF file. We had access to the DSMF files for 1989-1999 and, thus, the last known possible date of death would be December 31, 1999. Death dates prior to January 1, 1989 for the 1981-82 and 1986-87 samples were obtained from data previously compiled by Skonovd and Haapanen (2000). In this appendix, we describe the process used to obtain the dates of deaths for the deceased CYA wards in the three samples; this allowed for the addition of new, more recent death data for the 1981-82 and 1986-87 samples (i.e., recorded deaths through 1999), as well as enabling the collection of mortality data through December 31, 1999 for the 1991-92 release sample.

Due to the fact that DHS uses Social Security Numbers (SSN) as the "primary key" variable in their DSMF data files, and we did not have access to the SSN of the wards in the samples, the retrieval of the DHS records corresponding to the deceased CYA wards was, to say the least, a challenging task. The process of matching records

from the three samples to a record in the DSMF files occurred in seven steps described below and graphically depicted in Figure B.1. Cases only entered a subsequent step of the process upon successful completion of each prior step; the exception to this rule is Step (6).

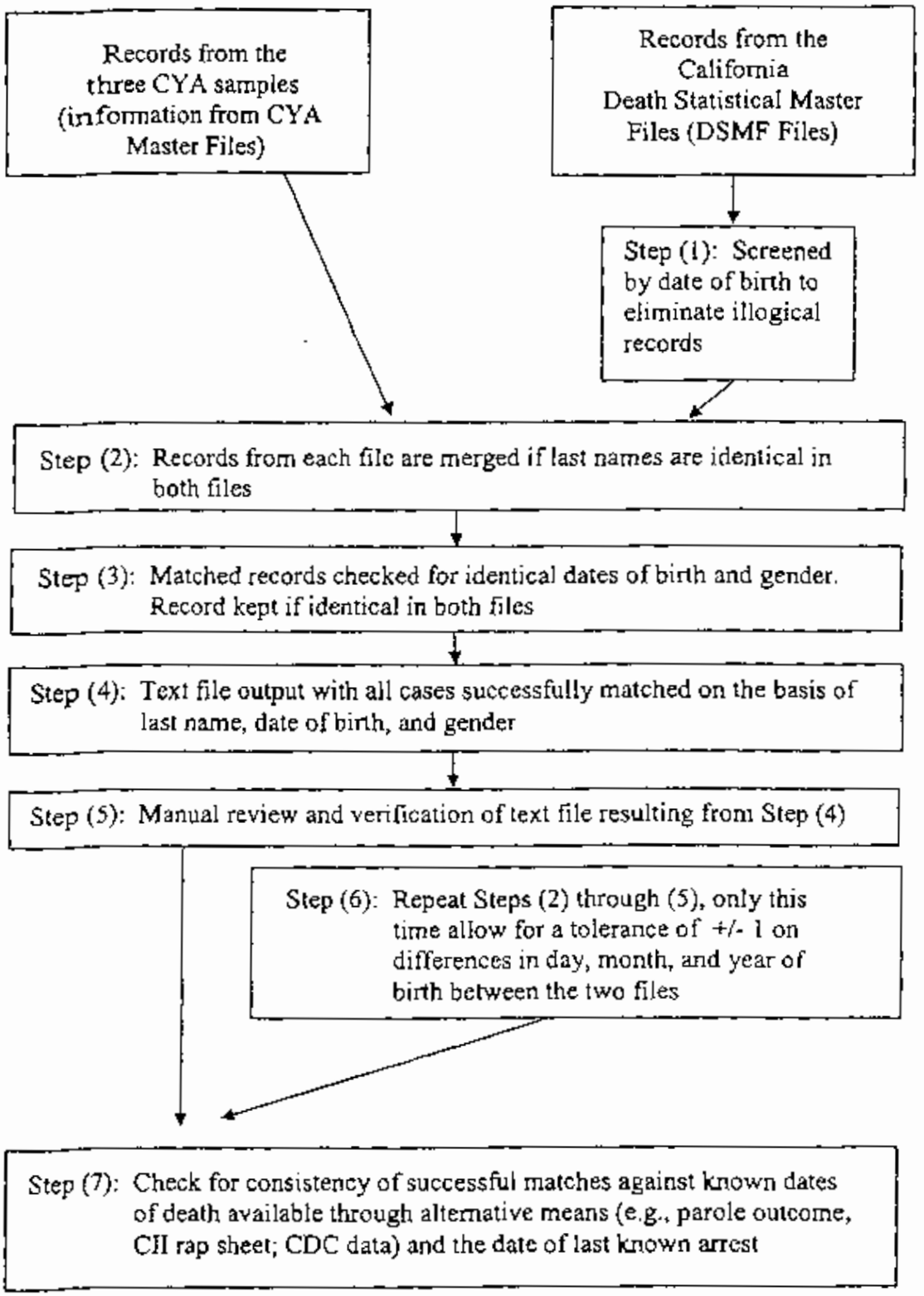
Step (1)

Due to the fact there were more than 200,000 deaths in California during each statistical year examined here, an initial record elimination step was undertaken in order to avoid both exceeding the computer system's hardware limits (e.g., running out memory and/or hard drive space) and an excessive amount of computer analysis time. This step involved eliminating significant portions of records in each DSMF file that were logically impossible matches. The DSMF files were initially screened to filter out all records that could not possibly have been a match due to the date of birth of the decedent. The earliest birth date year in any of the three samples was 1956 and the latest birth date was 1978, and therefore records with birth dates outside the range of 1956-1978 were initially excluded as potential matches.

Step (2)

The next step involved the remaining pool of potential matches left in the DSMF files after Step (1). In the second step, cases were next matched via a many-to-one match on the basis of the last names (with all letters treated as capital letters) in the DSMF file and the name the ward was admitted into the CYA under (i.e., last name from the CYA

Figure B.1. Graphical Representation of the Mortality Data Retrieval Process



master file). For each last name of the CYA wards, the cases were "matched" or "joined" to as many records in the DSMF files that shared the same last name. Thus, if the ward's last name was Smith and there were 100 records in the DSMF files with the last name of Smith, that one case would have been matched to 100 cases in the DSMF files.

Step (3)

Upon each established match of last name, the cases were then screened according to whether the date of birth and gender were exact matches in both the DSMF file and in the file containing the information from the CYA master file. Cases with different dates of birth and/or different gender were dropped at this point.

Step (4)

Upon a successful match of last name, date of birth, and gender, a text file was written-out containing the full names (first, middle and last names) and recorded ethnicity from both the CYA master files and the DSMF files.

Step (5)

The text file resulting from Step (4) was then manually reviewed and successful matches were determined on the basis of a comparison of first name, middle name, last name, and ethnicity. Although probabilistic matching methods and "sounds like" algorithms were implemented in an attempt to outsource the manual review to a computer algorithm, there simply was not probability cutoff point that reliably generated links between the files (we were not totally "blind" in this process since we had prior

information on the dates of death for over 30 wards in the 1991-92 sample that died while on parole). The matches had higher success rates when they were completed in the labor-intensive manual method of a detailed examination of the name components from each file, and therefore that method was used. Successful matches at the end of this step were then extracted from the text file and entered in Step (7).

Step (6)

In this step, Steps (2) through (5) were repeated, only this time we allowed some "tolerance" around the match of the dates of birth in the two files. More specifically, we allowed the day, month, and year of birth to vary +/- 1 in order to catch possible key entry errors in the dates of birth dates. Step (1) was altered accordingly. This step resulted in an additional 6 matches, all of which were manually reviewed and verified.

Step (7)

At this point, a series of different data checks were completed on the matched death records. First, all cases considered to be successful "matches" up to this point, including the cases with dates of death prior to 1990 that were collected by Skonovd and Happanen (2000), were checked against the last known arrest date for each case to assert that the last known arrest date occurred prior in time to the recorded date of death (after all, it's kind of difficult to be arrested when you're supposedly deceased). For the new death data (1990-1999), no case that was determined to be a successful match between the sample cases and the DHS death data was found to have a recorded arrest event after the date of death, although most of them had arrests prior to their deaths. For the prior

death data, however, 1 case was found to have had several arrests after the previously matched date of death. This case was removed as a mortality case. Second, homicides of known criminal offenders are often reported to the California Department of Justice and a record is attached to their CII rap sheet indicating that they were deceased (as a result of homicide). We checked all of the successful matches against the "death records" in the CII files and found that we had successfully retrieved the DHS death records for all but 4 of the cases with CII death records (n=111). For these 4 cases, it became quite clear how the process had failed; they all had different last names in the DHS death data, names that included a derivative form of the name recorded in the CYA master files (e.g., Smith, Smithfield; Jones, Joneston). Since the date of death was known from the CII records, the DHS records for these cases were manually retrieved by reviewing the DHS data file for the given death date and then finding the record that pertained to the case (which was how we discovered the "derivative name" reason for why the cases had not originally been retrieved in the second step). Finally, we compiled a list of cases known to have died on parole while we were coding the 1991-92 sample, and a comparison of the list of known fatalities matched against our list of "successful matches" produced a 100% match. That is, for all of the 1991-92 cases that we knew had died prior to completing parole or shortly thereafter (n=39), Steps (1) – (6) produced the DHS death record for all 39 of those cases. This lends credibility to our method of matching the records between the list of CYA cases and the fatality cases in the DHS DSMF files.

Limitations

Although we have great confidence in our method of matching cases between the two files (i.e., finding the CYA cases in the DHS death data), the resulting mortality file for the CYA cases studied here most certainly is an undercount of mortality in these samples for several reasons. First of all, the resulting mortality file only includes deaths that occurred in the state of California, and thus any deaths occurring outside of the state of California are not included here. We simply did not have access to mortality data outside of California. Secondly, although we made every effort possible to make sure we had matched all cases present in both files, there is a chance we did miss some cases due to either name changes (which is less problematic for the male wards than the female wards not included in this study) and incorrectly entered dates of birth that escaped our method of detecting cases that slipped through the initial matching process.

APPENDIX C

PERCENTAGE OF SAMPLES "AT RISK" OVER THE AGE DISTRIBUTION

Table C.1. Percentage of Cases "At Risk" of Arrest at Each Age, by Sample

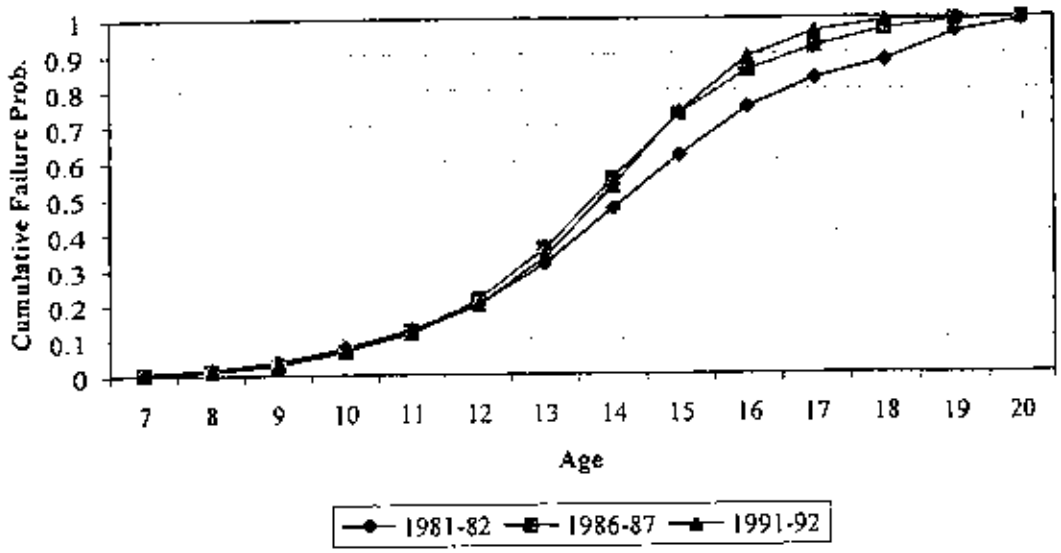
	Sample					
	1981-82		1986-87		1991-92	
Age	N	%	N	%	N	%
7	1989	100.00	1443	100.00	1434	100.00
8	1989	100.00	1443	100.00	1434	100.00
9	1989	100.00	1443	100.00	1434	100.00
10	1989	100.00	1443	100.00	1434	100.00
11	1989	100.00	1443	100.00	1434	100.00
12	1989	100.00	1443	100.00	1434	100.00
13	1989	100.00	1443	100.00	1434	100.00
14	1984	99.75	1440	99.79	1426	99.44
15	1965	98.79	1428	98.96	1405	97.98
16	1941	97.59	1372	95.08	1317	91.84
17	1905	95.78	1274	88.29	1193	83.19
18	1850	93.01	1175	81.43	1053	73.43
19	1863	93.67	1241	86.00	1197	83.47
20	1883	94.67	1332	92.31	1323	92.26
21	1926	96.83	1397	96.81	1353	94.35
22	1963	98.69	1401	97.09	1368	95.40
23	1957	98.39	1407	97.51	1381	96.30
24	1955	98.29	1405	97.37	1374	95.82
25	1948	97.94	1407	97.51	1333	92.96
26	1934	97.23	1394	96.60	1236	86.19
27	1927	96.88	1387	96.12	999	69.67
28	1917	96.38	1378	95.50	656	45.75
29	1904	95.73	1366	94.66	338	23.57
30	1898	95.42	1339	92.79	197	13.74
31	1891	95.07	1261	87.39	126	8.79
32	1879	94.47	1049	72.70	62	4.32
33	1867	93.87	751	52.04	10	0.70
34	1841	92.56	474	32.85	---	---
35	1739	87.43	263	18.23	---	---
36	1553	78.08	184	12.75	---	---
37	1243	62.49	113	7.83	---	---
38	896	45.05	51	3.53	---	---
39	529	26.60	---	---	---	---
40	252	12.67	---	---	---	---
41	89	4.47	---	---	---	---
42	25	1.26	---	---	---	---
43	6	0.30	---	---	---	---
Total Panel Observ.	60453		37390		29385	

APPENDIX D

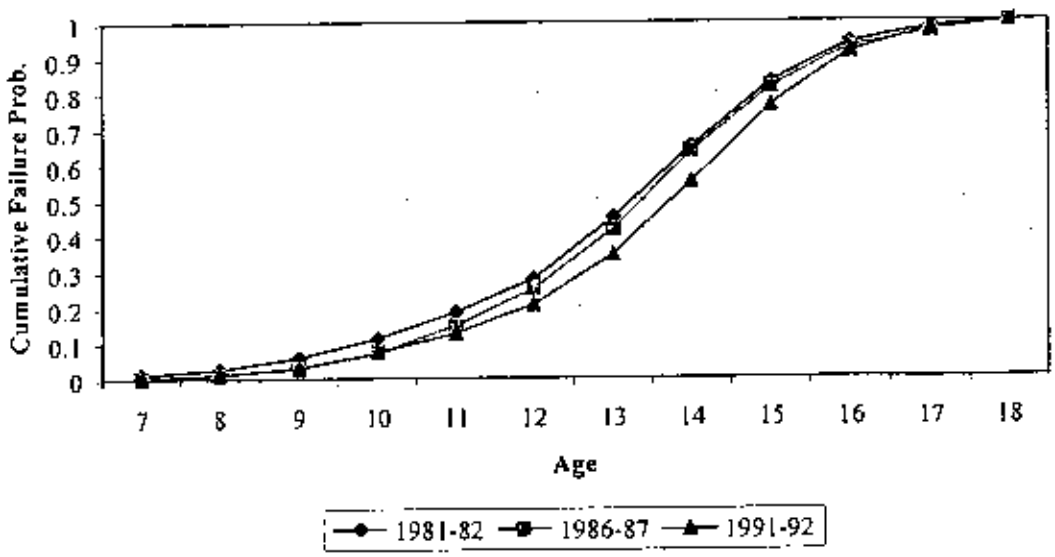
CUMULATIVE AGES OF ONSET OF FIRST CRIMINAL ARREST

Figure D.1. Cumulative Probaiblities of Age at First Criminal Arrest, by Sample

Panel A: All Wards



Panel B: Juvenile Court Commitments Only



APPENDIX E

MEANS OF BACKGROUND CHARACTERISTIC VARIABLES, BY LATENT CLASS

Table E.1. Means of Background Characteristic Variables, by Latent Class: 1981-82 Sample

	Latent Class					
Variable	LC1	LC2	LC3	LC4	LC5	LC6
White	0.34	0.29	0.31	0.46	0.48	0.39
Hispanic	0.24	0.27	0.28	0.25	0.26	0.20
African-American	0.40	0.41	0.40	0.25	0.22	0.40
Other	0.02	0.03	0.01	0.04	0.04	0.01
Family Violence	0.14	0.20	0.17	0.18	0.17	0.14
Parental Alc./Drug Dependence	0.22	0.31	0.31	0.25	0.31	0.24
Parental Criminality	0.19	0.25	0.25	0.07	0.24	0.17
Sibling Criminality	0.34	0.51	0.45	0.27	0.35	0.40
Lack of Supervision/Neglect	0.32	0.40	0.38	0.31	0.36	0.27
Ineffective Control	0.54	0.68	0.68	0.45	0.60	0.54
Physical Abuse	0.14	0.17	0.15	0.16	0.14	0.13
Sexual Abuse	0.01	0.01	0.00	0.02	0.02	0.02
Drug Abuse	0.68	0.69	0.70	0.68	0.67	0.69
Gang Member/Association	0.27	0.33	0.36	0.17	0.29	0.24
Previous Violent Behavior	0.78	0.85	0.83	0.72	0.77	0.81
School Dropout	0.61	0.49	0.50	0.63	0.44	0.61
DDMS Infacrtions	1.06	1.26	1.21	0.79	1.15	1.18
Deceased	0.06	0.05	0.07	0.06	0.22	0.05
Deceased	0.06	0.05	0.07	0.06	0.22	0.05

Note: All variables are binary variables (except for the number of DDMS infractions) and the means represent the percentage of cases coded as 1 (which indicates the presence of the characteristic).

Table E.2. Means of Background Characteristic Variables, by Latent Class: 1986-87 Sample

	Latent Class					
Variable	LC1	LC2	LC3	LC4	LC5	LC6
White	0.28	0.28	0.21	0.33	0.29	0.29
Hispanic	0.29	0.32	0.30	0.40	0.35	0.36
African-American	0.41	0.38	0.47	0.22	0.31	0.33
Other	0.01	0.02	0.02	0.05	0.05	0.03
Family Violence	0.20	0.16	0.21	0.18	0.17	0.17
Parental Alc./Drug Dependence	0.33	0.29	0.38	0.33	0.36	0.40
Parental Criminality	0.28	0.25	0.35	0.23	0.29	0.26
Sibling Criminality	0.38	0.37	0.46	0.32	0.31	0.43
Lack of Supervision/Neglect	0.41	0.36	0.49	0.44	0.51	0.46
Ineffective Control	0.74	0.75	0.84	0.75	0.79	0.76
Physical Abuse	0.24	0.16	0.21	0.21	0.16	0.18
Sexual Abuse	0.04	0.02	0.02	0.03	0.02	0.04
Drug Abuse	0.84	0.87	0.84	0.79	0.92	0.91
Gang Member/Association	0.58	0.50	0.55	0.49	0.49	0.39
Previous Violent Behavior	0.87	0.84	0.91	0.78	0.88	0.85
School Dropout	0.33	0.44	0.40	0.38	0.38	0.42
DDMS Infacrtions	2.12	1.65	2.68	2.04	3.17	1.56
Deceased	0.06	0.05	0.08	0.11	0.09	0.05

Note: All variables are binary variables (except for the number of DDMS infractions) and the means represent the percentage of cases coded as 1 (which indicates the presence of the characteristic).

Table E.3. Means of Background Characteristic Variables, by Latent Class: 1991-92 Sample

Variable	Latent Class					
	LC1	LC2	LC3	LC4	LC5	LC6
White	0.17	0.18	0.20	0.25	0.19	0.24
Hispanic	0.31	0.40	0.30	0.34	0.37	0.33
African-American	0.47	0.38	0.46	0.36	0.29	0.40
Other	0.05	0.05	0.04	0.05	0.15	0.03
Family Violence	0.19	0.26	0.24	0.28	0.17	0.20
Parental Alc./Drug Dependence	0.37	0.38	0.51	0.51	0.35	0.46
Parental Criminality	0.25	0.29	0.45	0.46	0.26	0.35
Sibling Criminality	0.41	0.38	0.46	0.52	0.40	0.43
Lack of Supervision/Neglect	0.59	0.61	0.74	0.77	0.54	0.76
Ineffective Control	0.77	0.84	0.93	0.97	0.86	0.92
Physical Abuse	0.19	0.22	0.23	0.27	0.22	0.19
Sexual Abuse	0.06	0.04	0.04	0.07	0.06	0.06
Drug Abuse	0.70	0.70	0.75	0.73	0.67	0.80
Gang Member/Association	0.70	0.77	0.74	0.75	0.77	0.75
Previous Violent Behavior	0.86	0.86	0.92	0.93	0.84	0.88
School Dropout	0.85	0.85	0.89	0.87	0.89	0.94
DDMS Infactions	1.95	1.81	2.27	2.22	1.48	2.80
Deceased	0.03	0.09	0.03	0.04	0.14	0.01

Note: All variables are binary variables (except for the number of DDMS infractions) and the means represent the percentage of cases coded as 1 (which indicates the presence of the characteristic).

APPENDIX F

OBSERVED AND PREDICTED ARREST TRAJECTORIES, BY LATENT CLASS

Figure F.1. Observed and Predicted Trajectories of Each Latent Class: 1981-82 Sample

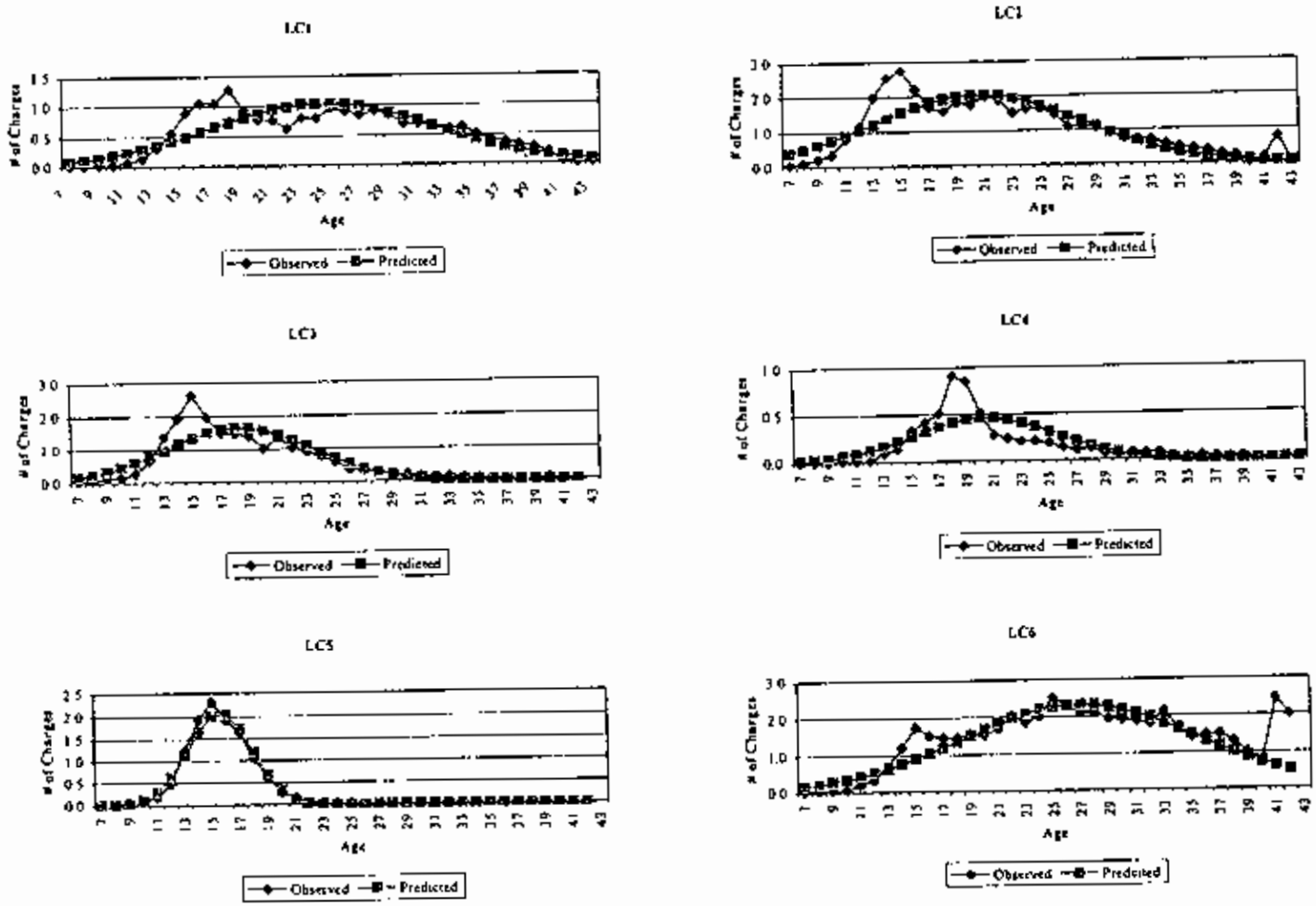


Figure F.2. Observed and Predicted Trajectories of Each Latent Class: 1986-87 Sample

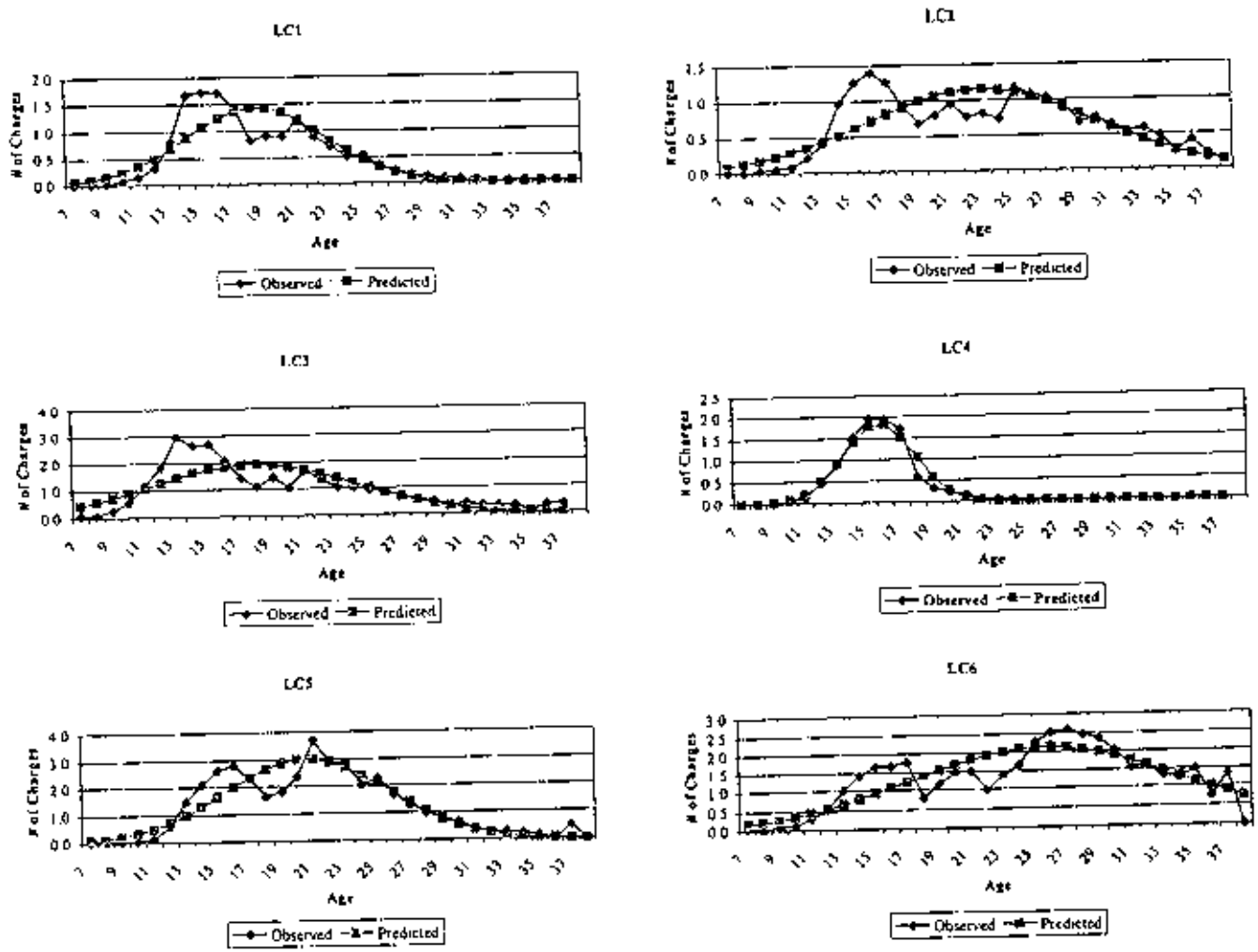
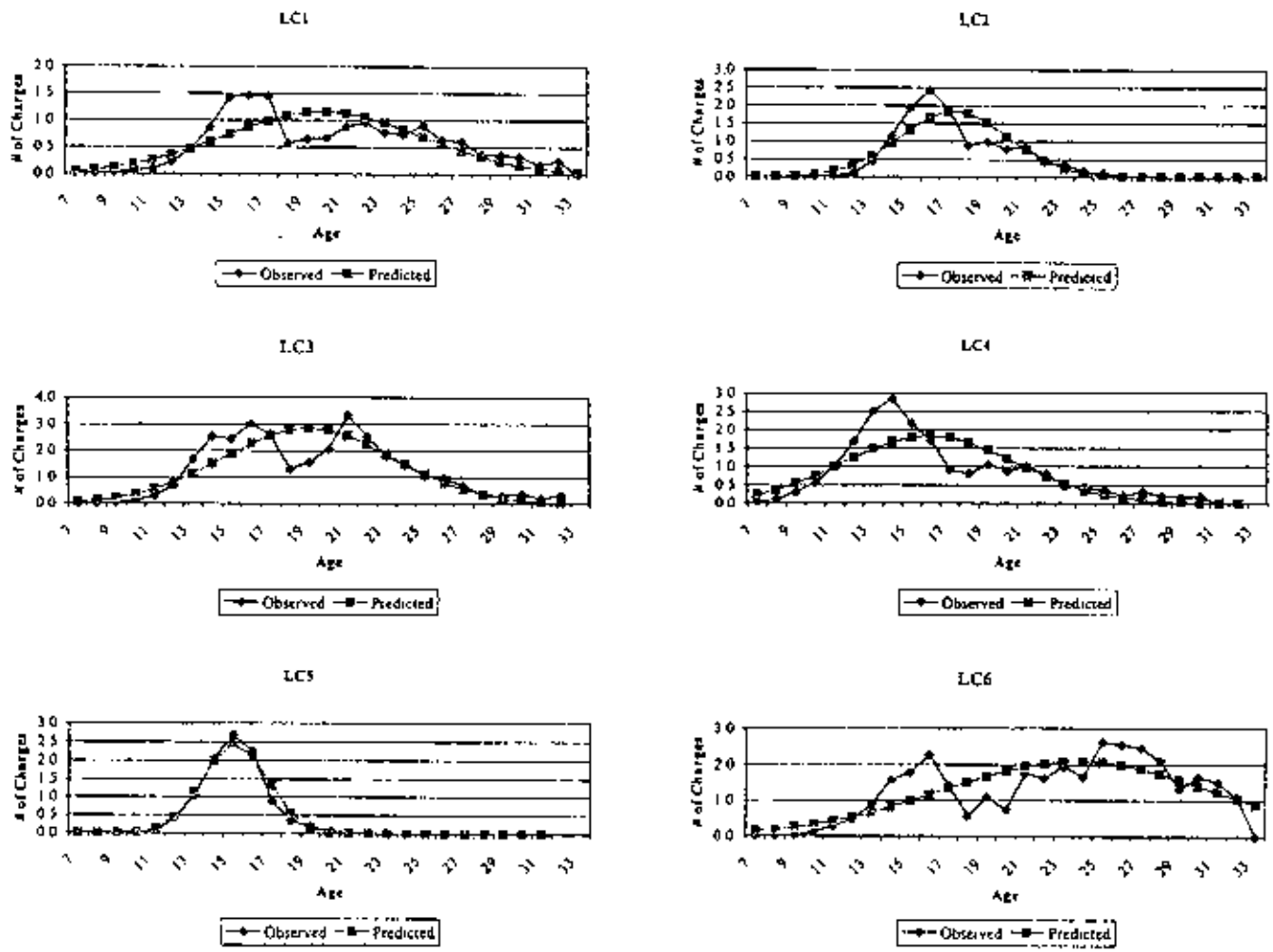


Figure F.3. Observed and Predicted Trajectories of Each Latent Class: 1991-92 Sample



REFERENCES

Agnew, Robert. 1992. "Foundation for a General Strain Theory of Crime and Delinquency." *Criminology* 30:47-87.

Aguilar, Benjamin, L. Alan Sroufe, Byron Egeland, and Elizabeth Carlson. 2000. "Distinguishing the Early-Onset/Persistent and Adolscence-Onset Antisocial Behavior Types: From Birth to 16 Years." *Development and Psychopathology* 12:109-32.

Akers, Ronald. 1985. *Deviant Behavior: A Social Learning Approach*. Belmont, CA: Wadsworth.

———. 1991. "Self-Control as a General Theory of Crime." *Journal of Quantitative Criminology* 7:201-11.

Alarcon, Francisco J. 1994."Verbal Testimony To The Little Hoover Commission." (March 17). Sacramento: CA. Department of the Youth Authority

American Law Institute. 1940. *Youth Authority Correction Act: Official Draft*. Washington, D.C.: American Law Institute.

Ayers, Charles D., James Herbert Williams, J. David Hawkins, Peggy L. Peterson, Richard F. Catalano, and Robert D. Abbott. 1999. "Assessing Correlates of Onset, Escalation, Deescalation, and Desistance of Delinquent Behavior." *Journal of Quantitative Criminology* 15:277-306.

Baltes, P. and J. Nesselroade. 1984. "Paradigm Lost and Paradigm Regained: Critique of Dannefer's Portrayal of Life-Span Developmental Psychology." *American Sociological Review* 49:841-46.

Barlow, Hugh. 1991. "Explaining Crimes and Analogous Acts, or the Unrestrained Will Grab at Pleasure Whenever They Can." *Journal of Criminal Law and Criminology* 82:229-42.

Bates, G. and J. Neyman. 1952. "Contributions to the Theory of Accident Proneness. II: True or False Contagion." *University of California Publications in Statistics* 1:255-75.

Beccaria, Cesare. 1963. *On Crimes and Punishment*. Translated by Henry Paolucci. Indianapolis, IN: Bobbs-Merrill.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76:169-217.

Benson, Michael L. 2002. *Crime and the Life Course*. Los Angeles: Roxbury Publishing Company.

Bernard, Thomas J. and R. Richard Ritti. 1991. "The Philadelphia Birth Cohort and Selective Incapacitation." *Journal of Research in Crime and Delinquency* 28:35-54.

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

———. 1987. "Characterizing Criminal Careers." *Science* 237:985-91.

Blumstein, A., J. Cohen, and D. Farrington. 1988a. "Criminal Career Research: Its Value for Criminology." *Criminology* 26:1-36.

———. 1988b. "Longitudinal and Criminal Career Research: Further Clarifications." *Criminology* 26:57-74.

Blumstein, A., J. Cohen, J. Roth, and C. A. Visher. 1986. *Criminal Careers and "Career Criminal."* Washington, D.C.: National Academy Press.

Blumstein, Alfred and Allen J. Beck. 1999. "Population Growth in U.S. Prisons, 1980-1996." Pp. 17-61 in *Crime and Justice: A Review of Research*, edited by M. Tonry and J. Petersilia. Chicago: University of Chicago Press.

Bolen, Jane. 1972. "The California Youth Authority 1941-1971: Structure, Policies and Practices," School of Social Work, University of Southern California.

Brame, Robert, Shawn Bushway, and Raymond Paternoster. 1999. "On the Use of Panel Research Designs and Random Effects Models to Investigate Static and Dynamic Theories of Criminal Offending." *Criminology* 37:599-641.

Brame, Robert, Raymond Paternoster, Paul Mazerolle, and Alex Piquero. 1998. "Testing for the Equality of Maximum-Likelihood Regression Coefficients Between Two Independent Equations." *Journal of Quantitative Criminology* 14:245-61.

Britt, C. L. 1992. "Constancy and Change in the U.S. Age Distribution of Crime: A Test of the "Invariance Hypothesis"." *Journal of Quantitative Criminology* 8:175-87.

Bureau of Justice Statistics. 1995. *Prisoners in 1994*. Washington, D.C.: U.S. Government Printing Office.

———. 2000. *Sourcebook of Criminal Justice Statistics*. Washington D.C.: U.S. Government Printing Office.

———. 2001. *Prisoners in 2000*. Washington, D.C.: U.S. Government Printing Office.

Burton, Paul, Lyle Gurrin, and Peter Sly. 1998. "Extending the Simple Linear Regression Model to Account for Correlated Responses: An Introduction to Generalized Estimating Equations and Multi-Level Mixed Modelling." *Statistics in Medicine* 17:1261-91.

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* 39:491-515.

Bushway, Shawn, Robert Brame, and Raymond Paternoster. 1999. "Assessing Stability and Change in Criminal Offending: A Comparison of Random Effects, Semiparametric, and Fixed Effects Modeling Strategies." *Journal of Quantitative Criminology* 15:23-61.

California Department of Justice. 1999. *Crime and Delinquency in California, 1998*. Sacramento: California Department of Justice.

———. 2000. *Crime and Delinquency in California, 1999*. Sacramento, CA: California Department of Justice.

———. 2001. *Crime and Delinquency in California, 2000*. Sacramento, CA: California Department of Justice.

California Youth Authority. 1971. *An Analysis of Institution Lengths of Stay by Selected Variables*. Sacramento, CA: Department of the Youth Authority.

———. 1981. *Analysis of Institutional and Parole Lengths of Stay, 1980*. Sacramento, CA: Department of the Youth Authority.

———. 1991. *Length of Stay of Youth Authority Wards Institutionalized and on Parole During 1990*. Sacramento, CA: Department of the Youth Authority.

———. 1999. *Length of Stay of Youth Authority Wards*. Sacramento, CA: Department of the Youth Authority.

Cameron, A. Colin and Pravin K. Trivedi. 1986. "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators." *Journal of Applied Econometrics* 1:29-53.

———. 1998. *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Canela-Cacho, Jose, Alfred Blumstein, and Jacqueline Cohen. 1997. "Relationship Between the Offending Frequency of Imprisoned and Free Offenders." *Criminology* 35:133-71.

Caplow, Theodore and Jonathon Simon. 1999. "Understanding Prison Policy and Population Trends." Pp. 63-120 in *Crime and Justice: A Review of Research*, edited by M. Tonry and J. Petersilia. Chicago: University of Chicago Press.

Caspi, A., D. Lynam, T.E. Moffitt, and P.A. Silva. 1993. "Unraveling Girls' Delinquency: Biological, Dispositional, and Contextual Contributions to Adolescent Misbehavior." *Developmental Psychology* 29:19-30.

Caspi, Avshalom and Terrie E. Moffitt. 1992. "The Continuity of Maladaptive Behavior: From Description to Understanding in the Study of Antisocial Behavior." In *Manual of Developmental Psychopathology*, edited by D. Cicchetti and D. Cohen. New York: Wiley.

———. 1993. "When Do Individual Differences Matter? A Paradoxical Theory of Personality Coherence." *Psychological Inquiry* 4:247-71.

Cernkovich, Stephen A. and Peggy C. Giordano. 2001. "Stability and Change in Antisocial Behavior: The Transition from Adolescence to Early Adulthood." *Criminology* 39:371-410.

Cernkovich, Stephen A., Peggy C. Giordano, and M.D. Pugh. 1985. "Chronic Offenders: The Missing Cases in Self-Report Delinquency Research." *Journal of Criminal Law and Criminology* 76:705-32.

Chung, Ick-Joong, Karl G. Hill, J. David Hawkins, Lewayne D. Gilchrist, and Daniel S. Nagin. 2002. "Childhood Predictors of Offense Trajectories." *Journal of Research in Crime and Delinquency* 39:60-90.

Cline, H.F. 1980. "Criminal Behavior Over the Life Span." Pp. 641-74 in *Constancy and Change in Human Development*, edited by O. Brim and J. Kagan. Cambridge, MA: Harvard University Press.

Clogg, Clifford C., Eva Petkova, and Adamantios Haritou. 1995. "Statistical Methods for Comparing Regression Coefficients Between Models." *American Journal of Sociology* 100:1261-93.

Cohen, Lawrence E. 1987. "Throwing Down the Gauntlet: A Challenge to the Relevance of Sociology for the Etiology of Criminal Behavior." *Contemporary Sociology* 16:202-05.

Cohen, Lawrence E. and Kenneth C. Land. 1987. "Age Structure and Crime: Symmetry Versus Asymmetry and the Projection of Crime Rates Through the 1990s." *American Sociological Review* 52:170-83.

Cohen, Lawrence E. and Bryan J. Vila. 1996. "Self-Control and Social-Control: An Exposition of the Gottfredson-Hirschi/Sampson-Laub Debate." *Studies on Crime and Crime Prevention* 5:125-50.

Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94:95-120.

———. 1990. *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Collin, A. Colin and Pravin K. Trivedi. 1996. "Count Data Models for Financial Data." In *Handbook of Statistics, Volume 14, Statistical Methods in Finance*, edited by G. Maddala and C. Rao. Amsterdam: North-Holland.

Dannefer, Dale. 1984. "Adult Development and Social Theory: A Paradigmatic Reappraisal." *American Sociological Review* 49:100-16.

———. 1987. "Aging as Intracohort Differentiation: Accentuation, the Matthew Effect, and the Life Course." *Sociological Forum* 2:211-36.

Dean, Charles W., Robert Brame, and Alex R. Piquero. 1996. "Criminal Propensities, Discrete Groups of Offenders, and Persistence in Crime." *Criminology* 34:547-74.

Dean, Charmaine B. 1998. "Overdispersion." Pp. 3226-32 in *Encyclopedia of Biostatistics*, edited by P. Armitage and T. Colton. Chichester, NY: John Wiley & Sons.

Dempster, A.P., N.M. Laird, and D.B. Rubin. 1977. "Maximum Likelihood Estimation from Incomplete Data Via the EM Algorithm (with Discussion)." *Journal of the Royal Statistical Society, Series B*. 39:1-38.

DeMuro, Paul, Anne DeMuro, and Steve Lerner. 1988. *Reforming the California Youth Authority*. Bolinas, CA: Common Knowledge Press.

Department of the Youth Authority. 1994. *Written Testimony for the Little Hoover Commission: Mission and Values*. Sacramento, CA: Department of the Youth Authority.

Department of the Youth Authority. 1994. *Written Testimony for the Little Hoover Commission: Perspectives and Issues on Delinquency Prevention*. Sacramento, CA: Department of the Youth Authority.

Dobrin, Adam. 2001. "The Risk of Offending on Homicide Victimization: A Case Control Study." *Journal of Research in Crime and Delinquency* 38:154-73.

Donziger, Steven R., editor. 1996. *The Real War on Crime*. New York: Harper Collins.

Durkheim, Emilie. [1897] 1951. *Suicide*. Translated by J. Spaulding and G. Simpson. Reprint, New York: Free Press.

D'Unger, Amy V., Kenneth C. Land, and Patricia L. McCall. Forthcoming. "Gender Differences in Age Patterns of Delinquent/Criminal Careers: Results from Poisson Latent Class Analysis of the Philadelphia Cohort Study." *Journal of Quantitative Criminology*.

D'Unger, Amy V., Kenneth C. Land, Patricia L. McCall, and Daniel S. Nagin. 1998. "How Many Latent Classes of Delinquent/Criminal Careers? Results from Mixed Poisson Regression Analyses." *American Journal of Sociology* 103:1593-630.

Elder, Glen H., Jr. 1985. "Perspectives on the Life Course." Pp. 23-49 in *Life Course Dynamics*, edited by G. H. Elder, Jr. Ithaca: Cornell University Press.

———. 1992. "The Life Course." Pp. 1120-30 in *The Encyclopedia of Sociology*, edited by E. F. Borgatta and M. L. Borgatta. New York: Macmillan.

———. 1996. "Human Lives in Changing Societies: Life Course and Developmental Insights." Pp. 31-62 in *Developmental Science*, edited by R. B. Cairns, G. H. J. Elder and E. J. Costello. Cambridge: Cambridge University Press.

Elder, Glen H., Jr and Angela M. O'Rand. 1995. "Adult Lives in a Changing Society." Pp. 452-75 in *Sociological Perspectives in Social Psychology*, edited by K. S. Cook, G. A. Fine and J. S. House. Boston: Allyn & Bacon.

Elliot, Delbert S. 1994. "Serious Violent Offenders: Onset, Developmental Course, and Termination -- the American Society of Criminology 1993 Presidential Address." *Criminology* 32:1-21.

Ezell, Michael E. and Lawrence E. Cohen. 1997. "Age, Crime, and Crime Control Policies: A Longitudinal Analysis of Youthful, Serious, Chronic Offenders with Implications for the "Three Strikes" Legislation." *Studies on Crime and Crime Prevention* 6:169-99.

Farrington, D. 1983. *Further Analyses of a Longitudinal Survey of Crime and Delinquency*. Final Report to the National Institute of Justice Washington D.C.: National Institute of Justice.

Farrington, David P. 1986. "Age and Crime." Pp. 189-250 in *Crime and Justice: An Annual Review of Research*, edited by M. Tonry and N. Morris. Chicago: University of Chicago Press.

———. 1989. "Self-Reported and Official Offending from Adolescence to Adulthood." Pp. 399-423 in *Cross-National Research in Self-Reported Crime and Delinquency*, edited by M. Klein. Dordrecht, the Netherlands: Kluwer.

———. 1997. "Human Development and Criminal Careers." Pp. 361-408 in *The Oxford of Criminology*, edited by M. Maguire, R. Morgan and R. Reiner. Oxford: Clarendon Press.

Farrington, David P., Loeber, Delbert S. Elliot, David Hawkins, Denise B. Kandel, Malcolm W. Klein, Joan McCord, David C. Rowe, and Richard E. Tremblay. 1991. "Advancing Knowledge About the Onset of Delinquency and Crime." Pp. 283-342 in *Advances in Clinical Child Psychology, Volume 13*. New York: Plenum.

Fergusson, David M., L. John Horwood, and Daniel S. Nagin. 2000. "Offending Trajectories in a New Zealand Birth Cohort." *Criminology* 38:525-51.

Fischer, M., J.E. Rolf, J.E. Hasazi, and L. Cummings. 1984. "Follow-up of a Pre-School Epidemiological Sample: Cross Age Continuities and Predictions of Later Adjustment with Internalizing and Externalizing Dimensions of Behavior." *Child Development* 55:137-50.

Ge, Xiaojia, M. Brent Donnellan, and Ernst Wenk. 2001. "The Development of Persistent Criminal Offending in Males." *Criminal Justice and Behavior* 28:731-55.

Geerken, Michael R. 1994. "Rap Sheets in Criminological Research: Considerations and Caveats." *Journal of Quantitative Criminology* 10:3-21.

Ghosh, J. K. and P. K. Sen. 1985. "On the Asymptotic Performance of the Log-Likelihood Ratio Statistic for the Mixture Model and Related Results." Pp. 789-806 in *Proceedings of the Berkeley Conference in Honor of Jerry Neyman and Jack Kiefer, Volume 2*, edited by L. M. LeCam and P. A. Olshen. Monterey, CA: Wadsworth.

Gibbons, Don C. 1982. *Society, Crime, and Criminal Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Glueck, Sheldon and Eleanor Glueck. 1950. *Unraveling Juvenile Delinquency*. Cambridge, MA: Harvard University Press.

———. 1968. *Delinquents and Nondelinquents in Perspective*. Cambridge, MA: Harvard University Press.

Goodman, L.A. 1974. "Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models." *Biometrika* 61:215-31.

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* 24:213-33.

———. 1987a. "The Methodological Adequacy of Longitudinal Research on Crime." *Criminology* 25:581-614.

———. 1987b. *Positive Criminology*. Newbury Park, CA: Sage Publications.

———. 1988. "Science, Public Policy, and the Career Paradigm." *Criminology* 26:37-55.

———. 1990. *A General Theory of Crime*. Stanford: Stanford University Press.

Gourieroux, C., A. Monfort, and A. Trognon. 1984a. "Pseudo Maximum Likelihood Methods: Theory." *Econometrica* 52:681-700.

———. 1984b. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica* 52:701-29.

Gove, Walter R. 1985. "The Effect of Age and Gender on Deviant Behavior: A Biopsychological Perspective." Pp. 115-44 in *Gender and the Life Course*, edited by A. S. Rossi. New York: Aldine de Gruyter.

Grasmick, Harold G, Charles R. Tittle, Robert J. Bursik, Jr., and Bruce J. Arneklev. 1993. "Testing the Core Empirical Implications of Gottfredson and Hirschi's General Theory of Crime." *Journal of Research in Crime and Delinquency* 30:5-29.

Greenberg, David F. 1977. "Delinquency and the Age Structure of Society." *Contemporary Crises* 1:189-224.

———. 1985. "Age, Crime, and Social Explanation." *American Journal of Sociology* 91:859-66.

Greene, William H. 1998. *LIMDEP, Version 7.0*. Belport, NY: Econometric Software.

———. 2000. *Econometric Analysis*, 4th edition. Upper Saddle River, NJ: Prentice-Hall, Inc.

Greenwood, M. and G.U. Yule. 1920. "An Inquiry Into the Nature and Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks or of Repeated Accidents." *Journal of the Royal Statistical Society* 83:255-79.

Gutierrez, Roberto G., Shana Carter, and David M. Drukker. 2001. "On Boundary-Value Likelihood-Ratio Tests." *Stata Technical Bulletin* 60:15-20.

Haapanen, Rudy. 1990. *Selective Incapacitation and the Serious Offender: A Longitudinal Study of Criminal Career Patterns*. New York: Springer-Verlag.

Haapanen, Rudy and Wes Ingram. 2000. *California Youth Authority Mental Health and Substance Abuse Treatment Needs Assessment: Description and Preliminary Findings*. Sacramento, CA: California Youth Authority.

Haapanen, Rudy and Lee Britton. 2002. "Drug Testing for Youthful Offenders on Parole: An Experimental Evaluation." *Criminology and Public Policy* 1:217-44.

Haberman, S.J. 1988. "A Stabilized Newton-Raphson Algorithm for Log-Linear Models for Frequency Tables Derived by Indirect Observation." Pp. 193-211 in *Sociological Methodology 1988*, edited by C. Clogg. San Francisco: Jossey-Bass.

Hagan, John and Alberto Palloni. 1988. "Crimes as Social Events in the Life Course: Reconceiving a Criminological Controversy." *Criminology* 26:87-100.

Hamerle, Alfred and Gerd Ronning. 1995. "Panel Analysis for Qualitative Variables." Pp. 401-52 in *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, edited by G. Arminger, C. C. Clogg and M. E. Sobel. New York: Plenum Press.

Hardin, James and Joseph Hilbe. 2001. *Generalized Linear Models and Extensions*. College Station, TX: Stata Corporation.

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica* 52:909-39.

Heckman, James J. 1981a. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Stochastic Process." In *Structural Analysis of Discrete Data with Econometric Applications*, edited by C. F. Manski and D. McFadden. Cambridge, MA: MIT Press.

———. 1981b. "Heterogeneity and State Dependence." Pp. 91-139 in *Studies in Labor Markets*, edited by S. Rosen. Chicago: University of Chicago Press.

Heckman, James J. and J. George Borjas. 1980. "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence." *Economica* 47:247-83.

Heckman, James J. and Burton Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Analysis." *Econometrica* 52:909-38.

Hindelang, M. J., Travis Hirschi, and J. G. Weis. 1979. "Correlates of Delinquency: The Illusion of Discrepancy Between Self-Report and Official Measures." *American Sociological Review* 44(995-1014).

Hindelang, M.J., T. Hirschi, and J.G. Weis. 1981. *Measuring Delinquency*. Beverly Hills, CA: Sage Publications.

Hirschi, Travis. 1969. *Causes of Delinquency*. Berkeley, CA: University of California Press.

———. 1995. "The Family." Pp. 121-40 in *Crime*, edited by J. Q. Wilson and J. Petersilia. San Francisco: Institute for Contemporary Studies Press.

Hirschi, Travis and Michael R. Gottfredson. 1983. "Age and the Explanation of Crime." *American Journal of Sociology* 89:552-84.

———. 1985. "Age and Crime, Logic and Scholarship: Comment on Greenberg." *American Journal of Sociology* 89:552-84.

———. 1988. "Toward a General Theory of Crime." Pp. 8-26 in *Explaining Criminal Behavior*, edited by W. Buikhuisen and S. Mednick. Leiden: E.J. Brill.

———. 1993. "Commentary: Testing the General Theory of Crime." *Journal of Research in Crime and Delinquency* 30:52-53.

Horney, Julie, D. Wayne Osgood, and Ineke Haen 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.

Hsiao, Cheng. 1986. *Analysis of Panel Data*. Cambridge, UK: Cambridge University Press.

Huizinga, D. and D.S. Elliot. 1986. "Reassessing the Reliability and Validity of Self-Report Measures." *Journal of Quantitative Criminology* 2:293-327.

Humes, Edward. 1996. *No Matter How Loud I Shout*. New York: Touchstone.

Irwin, John and James Austin. 1997. *It's About Time America's Imprisonment Binge.* Belmont, CA: Wadsworth Publishing Company.

Jeffrey, C. 1972. "The Historical Development of Criminology." In *Pioneers in Criminology*, edited by H. Mannheim. Montclair: Patterson-Smith.

Kass, Robert E. and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90:773-95.

Keribin, C. 1997. "Consistent Estimation of the Order of Mixture Models (Working Paper #61)," Universite d'Evry-val d'Essonne, Laboratorie Analyse et Probabilite.

Kessing, Lars Vedal, Elisabeth Wreford Olsen, and Per Kragh Andersen. 1999. "Recurrence of Affective Disorder: Analysis with Frailty Models." *American Journal of Epidemiology* 149:404-11.

King, Gary. 1988. "Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model." *American Journal of Political Science* 32:838-63.

———. 1989. "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator." *American Journal of Political Science* 33:762-84.

Klein, Malcolm W. 1995. "Street Gang Cycles." Pp. 217-36 in *Crime*, edited by J. Q. Wilson and J. Petersilia. San Francisco: Institute for Contemporary Studies.

Klevens, Joanne, Ofelia Restrepo, Juanita Roca, and Adriana Martinez. 2000. "Comparison of Offending with Early- and Late-Starting Antisocial Behavior in Colombia." *Criminology* 44:194-203.

Krisberg, Barry. 1987. *California's Juvenile Justice System in Turmoil.* San Francisco, CA. National Council on Crime and Delinquency.

Lahey, Benjamin B., Sherryl H. Goodman, Irwin D. Waldman, Hector Bird, Glorisa Canino, Peter Jensen, Darrel Regier, Philip J. Leaf, Rachel Gordon, and Brooks Applegate. 1999. "Relation of Age of Onset to the Type and Severity of Child and Adolescent Conduct Problems." *Journal of Abnormal Child Psychology* 27:247-60.

Lambert, D. 1992. "Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing." *Technometrics* 34:1-14.

Land, Kenneth C. 1992. "Models of Criminal Careers: Some Suggestions for Moving Beyond the Current Debate." *Criminology* 30:149-55.

Land, Kenneth C., Patricia L. McCall, and Daniel S. Nagin. 1996. "A Comparison of Poisson, Negative Binomial, and Semiparametric Mixed Poisson Regression Models." *Sociological Methods & Research* 24:387-442.

Land, Kenneth C. and Daniel S. Nagin. 1996. "Micro-Models of Criminal Careers: A Synthesis of the Criminal Careers and Life Course Approaches Via Semiparametric Mixed Poisson Regression Models, with Empirical Applications." *Journal of Quantitative Criminology* 12:163-91.

Land, Kenneth C., Daniel S. Nagin, and Patricia L. McCall. 2001. "Discrete-Time Hazard Regression Models with Hidden Heterogeneity." *Sociological Methods and Research* 29:342-73.

Lattimore, Pamela K., Richard L. Linster, and John M. MacDonald. 1997. "Risk of Death Among Serious Young Offenders." *Journal of Research in Crime and Delinquency* 34:187-209.

Laub, John H., Nagin Daniel S., and Robert J. Sampson. 1998. "Trajectories of Change in Criminal Offending: Good Marriages and the Desistance Process." *American Sociological Review* 63:225-38.

Laub, John H. and Robert J. Sampson. 1988. "Unraveling Families and Delinquency: A Reanalysis of the Gluecks' Data." *Criminology* 26:355-79.

———. 1993. "Turning Points in the Life Course: Why Change Matters to the Study of Crime." *Criminology* 31:301-25.

———. 1998. "The Long-Term Reach of Adolescent Competence: Socioeconomic Achievement in the Lives of Disadvantaged Men." In *Competence and Character Through Life*, edited by A. Colby, J. James and D. Hart. Chicago: University of Chicago Press.

———. 2001. "Understanding Desistence from Crime." Pp. 1-69 in *Crime and Justice: An Annual Review of Research*, edited by M. Tonry. Chicago: University of Chicago Press.

Laub, John H., Robert J. Sampson, Ronald P. Corbett, Jr., and Jinny S. Smith. 1995. "The Public Policy Implications of a Life-Course Perspective on Crime." Pp. 91-106 in *Crime and Public Policy*, edited by H. D. Barlow. Boulder, CO: Westview Press.

Lauritsen, Janet L. 1998. "The Age-Crime Debate." *Social Forces* 77:127-55.

———. 1999. "Limitations in the Use of Longitudinal Self-Report Data: A Comment." *Criminology* 37:687-94.

LeBlanc, M. and R. Loeber. 1998. "Developmental Criminology Updated." Pp. 115-98 in *Crime and Justice: An Annual Review of Reserch*. Chicago: University of Chicago Press.

Legislative Analyst's Office. 1994. *Crime in California*. Sacramento, CA: State of California.

———. 1995. *Juvenile Crime*. Sacramento, CA.

———. 1996. *Analysis of the 1996-97 Budget*. Sacramento, CA.

———. 2001. *Analysis of the 2001-02 Budget Bill*. Sacramento, CA.

Lemert, E.M. 1972. *Human Deviance, Social Problems, and Social Control*. Englewood Cliffs, NJ: Prentice Hall.

Lerner, Steve. 1982. *The CYA Report: Conditions of Life at the California Youth Authority*. Bolinas, CA: Common Knowledge Press.

———. 1986. *Bodily Harm: The Pattern of Fear and Violence at the California Youth Authority*. Bolinas, CA: Common Knowledge Press.

———. 1988. *The Good News about Juvenile Justice*. Bolinas, CA: Common Knowledge Press.

Lindsey, J. K. 1995. *Modelling Frequency and Count Data*. Oxford: Clarendon Press.

Lindsey, J.K. 1993. *Models for Repeated Measurements*. Oxford: Clarendon Press.

Little Hoover Commission. 1994. *The Juvenile Crime Challenge: Making Prevention a Priority*. Sacramento, CA: State of California.

Loeber, R. and M. Le Blanc. 1990. "Toward a Developmental Criminology." In *Crime and Justice: A Review of Research*, edited by M. Tonry and N. Morris. Chicago: University of Chicago Press.

Loeber, Rolf. 1982. "The Stability of Antisocial and Delinquent Child Behavior: A Review." *Child Development* 53:1431-46.

Loeber, Rolf and David P. Farrington. 1998. "Conclusions and the Way Forward." Pp. 405-27 in *Serious and Violent Juvenile Offenders*, edited by R. Loeber and D. P. Farrington. Thousand Oaks, CA: Sage Publications.

Loeber, Rolf, David P. Farrington, and Daniel A. Waschbusch. 1998. "Serious and Violent Juvenile Offenders." Pp. 13-29 in *Serious and Violent Juvenile Offenders*, edited by R. Loeber and D. P. Farrington. Thousand Oaks, CA: Sage Publications.

Loeber, Rolf and Magda Stouthamer-Loeber. 1996. "The Development of Offending." *Criminal Justice and Behavior* 23:12-24.

Loeber and Stouthamer-Loeber. 1998. "The Development of Juvenile Aggression and Violence: Some Common Misconceptions and Controversies." *American Psychologist* 52:242-59.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications.

Maddala, G.S. 1987. "Limited Dependent Variable Models Using Panel Data." *Journal of Economic Resources* 22:307-38.

Maltz, Michael D. 1994. "Deviating from the Mean: The Declining Significance of Significance." *Journal of Research in Crime and Delinquency* 31:434-63.

Maxson, Cheryl L. 1999. "Gang Homicide a Review and Extension of the Literature." Pp. 239-54 in *Homicide: A Sourcebook of Social Research*, edited by Smith and Zahn. Thousand Oaks, CA: Sage Publications.

Mazerolle, Paul, Robert Brame, Ray Paternoster, Alex Piquero, and Charles Dean. 2000. "Onset Age, Persistence, and Offending Versatility: Comparisons Across Gender." *Criminology* 38:1143-72.

McCord, Joan. 1980. "Patterns of Deviance." Pp. 157-65 in *Human Functioning in Longitudinal Perspective*, edited by S. Sells, R. Crandall, M. Roff, J. S. Strauss and W. Pollin. Baltimore: Williams and Wilkins.

McCullagh, P. and J.A. Nelder. 1989. *Generalized Linear Models*. 2. London: Chapman & Hall.

McDermott, Joan. 1983. "The Serious Juvenile Offender Problems in Definition and Targeting." Pp. 67-90 in *Evaluating Juvenile Justice*. Beverly Hills, CA: Sage Publications, Inc.

McDermott, Shaun and Daniel S. Nagin. 2001. "Same or Different? Comparing Offender Groups and Covariates Over Time." *Sociological Methods and Research* 29:282-318.

Moffitt, Terrie E. 1993. "Adolescence-Limited and Life-Course-Persistent Antisocial Behavior: A Developmental Taxonomy." *Psychological Review* 100:674-701.

———. 1997. "Adolescence-Limited and Life-Course-Persistent Offending: A Complementary Pair of Developmental Theories." Pp. 11-54 in *Developmental Theories of Crime and Delinquency*, edited by T. P. Thornberry. Advances in Criminological Theory, vol. 7. New Brunswick: Transaction Publishers.

Mullahy, J. 1986. "Specification and Testing of Some Modified Count Data Models." *Journal of Econometrics* 33:341-65.

Nagin, Daniel S. 1999. "Analyzing Developmental Trajectories: A Semiparametric, Group-Based Approach." *Psychological Methods* 4:139-57.

Nagin, Daniel S. and David P. Farrington. 1992a. "The Stability of Criminal Potential from Childhood to Adulthood." *Criminology* 30:235-59.

———. 1992b. "The Onset and Persistence of Offending." *Criminology* 30:501-23.

Nagin, Daniel S., David P. Farrington, and Terrie E. Moffitt. 1995. "Life-Course Trajectories of Different Types of Offenders." *Criminology* 33:111-37.

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 Raymond Paternoster. 1991. "On the Relationship of Past to Future Participation in Delinquency." *Criminology* 29:163-89.

———. 1993. "Enduring Individual Differences and Rational Choice Theories of Crime." *Law & Society Review* 27:467-96.

———. 1994. "Personal Capital and Social Control: The Deterrence Implications of a Theory of Individual Differences in Criminal Offending." *Criminology* 32:581.

Nagin, Daniel and Raymond Paternoster. 2000. "Population Heterogeneity and State Dependence: State of the Evidence and Directions for Future Research." *Journal of Quantitative Criminology* 16:117-44.

Nagin, Daniel and Richard E. Tremblay. 1999. "Trajectories of Boys' Physical Aggression, Opposition, and Hyperactivity on the Path to Physically Violent and Nonviolent Juvenile Delinquency." *Child Development* 70:1181-96.

Nelder and Wedderburn. 1972. "Generalized Linear Models." *Journal of the Royal Statistical Society, Series B* 135:370-84.

Neyman, J. 1965. "Certain Chance Mechanisms Involving Discrete Distributions." Pp. 1-14 in *Classical and Contagious Discrete Distributions*, edited by G. N. Patil. Calcutta: Statistical Publishing House.

Paternoster, Raymond and Robert Brame. 1997. "Multiple Routes to Delinquency? A Test of Developmental and General Theories of Crime." *Criminology* 35:49-84.

Paternoster, Raymond, Robert Brame, and David P. Farrington. 2001. "On the Relationship Between Adolescent and Adult Conviction Frequencies." *Journal of Quantitative Criminology* 17:201-25.

Paternoster, Raymond, Robert Brame, Paul Mazerolle, and Alex Piquero. 1998. "Using the Correct Statistical Test for the Equality of Regression Coefficients." *Criminology* 36:859-66.

Paternoster, Raymond, Charles W. Dean, Alex Piquero, Mazerolle, and Robert Brame. 1997. "Generality, Continuity, and Change in Offending." *Journal of Quantitative Criminology* 13:231-66.

Patterson, Gerald R. 1995. "Orderly Change in a Stable World: The Antisocial Trait as a Chimera." Pp. 83-102 in *The Analysis of Change*, edited by J. M. Gottman. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Patterson, Gerald R., Barbara B. DeBaryshe, and Elizabeth Ramsey. 1989. " A Developmental Perspective on Antisocial Behavior." *American Psychologist* 44:329-35.

Patterson, Gerald R., John B. Reid, and Thomas J. Dishon. 1992. *Antisocial Boys*. Eugene, OR: Castilia Publishing Co.

Patterson, Gerald R. and Karen Yoeger. 1993. "Developmental Models for Delinquent Behavior." In *Mental Disorder and Crime*, edited by S. Hodgins. Newbury Park, CA: Sage.

Pickles, A. 1998. "Generalized Estimating Equations." Pp. 1626-36 in *Encyclopedia of Biostatistics*, edited by P. Armitage and T. Colton. Chichester, NY: John Wiley & Sons.

Piquero, Alex R. and Timothy Brezina. 2001. "Testing Moffitt's Account of Adolescence-Delimited Delinquency." *Criminology* 39:353-70.

Piquero, Alex R., Randall MacIntosh, and Matthew Hickman. 2002. "The Validity of a Self-Reported Delinquency Scale." *Sociological Methods and Research* 30:492-529.

Piquero, Alex, Alfred Blumstein, Robert Brame, Rudy Haapanen, Edward P. Mulvey, and Daniel S. Nagin. 2001. "Assessing the Impact of Exposure Time and Incapacitation on Longitudinal Trajectories of Criminal Offending." *Journal of Adolescent Research* 16:54-74.

Piquero, Alex, Raymond Paternoster, Paul Mazerolle, Robert Brame, and Charles W. Dean. 1999. "Onset Age and Offense Specialization." *Journal of Research in Crime and Delinquency* 36:275-99.

Powers, Daniel A. and Yu Xie. 2000. *Statistical Methods for Categorical Data Analysis*. San Diego: Academic Press.

Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." *Sociological Methodology* 25:111-64.

Riley, Mathilda White. 1986. "Overview and Highlights of a Sociological Perspective." Pp. 153-75 in *Human Development and the Life Course: Multidisciplinary Perspectives*, edited by A. Sorensen, F. Weinert and L. Sherrod. Hillsdale, NJ: Lawrence Erlbaum.

Robins, Lee N. 1966. *Deviant Children Grown up*. Baltimore: Williams & Wilkins.

———. 1978. "Sturdy Childhood Predictors of Adult Antisocial Behavior." *Psychological Medicine* 8:611-22.

Roeder, Kathryn, Kevin G. Lynch, and Daniel S. Nagin. 1999. "Modeling Uncertainty in Latent Class Membership: A Case Study in Criminology." *Journal of the American Statistical Association* 94:766-76.

Rutter, Michael and Marjorie Rutter. 1993. *Developing Minds: Challenge and Continuity Across the Life Span*. New York: Basic Books.

Sampson, Robert J. 1992. "A General Theory of Crime." *Social Forces* 71:545-46.

———. 1998. "Continuity and Discontinuity in Criminal Careers." *Social Forces* 76:1149-50.

———. 2000. "Whither the Sociological Study of Crime?" *American Sociological Review* 26:711-14.

Sampson, Robert J. and John H. Laub. 1990. "Crime and Deviance Over the Life Course: The Salience of Adult Social Bonds." *American Sociological Review* 55:609-27.

———. 1992. "Crime and Deviance in the Life Course." *Annual Review of Sociology* 18:63-84.

———. 1993. *Crime in the Making Pathways and Turning Points Through Life.* Cambridge, MA: Harvard University Press.

———. 1997. "A Life-Course Theory of Cumulative Disadvantage and the Stability of Delinquency." Pp. 133-61 in *Developmental Theories of Crime and Delinquency*, edited by T. P. Thornberry. Advances in Criminological Theory, vol. 7. New Brunswick: Transaction Publishers.

Self, S.G. and K. Y. Liang. 1987. "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood-Ratio Tests Under Nonstandard Conditions." *Journal of the American Statistical Association* 82:605-10.

Shannon, L. W., J. L. McKim, J. P. Curry, and L. J. Haffner. 1988. *Criminal Career Continuity: Its Social Context.* New York: Human Sciences Press.

Shannon, Lyle W. 1988. *Criminal Career Continuity.* New York: Human Sciences Press.

———. 1991. *Changing Patterns of Delinquency and Crime: A Longitudinal Study in Racine.* Boulder, CO: Westview Press.

Shavit, Yossi and Arye Rattner. 1988. "Age, Crime, and the Early Life Course." *American Journal of Sociology* 6:1457-70.

Skonovd, Norman. 1986. *Youth Authority Institutional Length of Stay and Recidivism: California Juvenile Arrest Rates and Guidelines for Parole Consideration Dates.* Sacramento, CA: Department of the Youth Authority.

Skonovd, Norman and Rudy Haapanen. 2000. *Predicting Parole Performance in the Era of Crack Cocaine.* Sacramento, CA: Deparment of the Youth Authority.

Snyder, H. 2002. *Juvenile Arrests 2000.* Washington, D.C.,: Office of Juvenile Justice and Delinquency Prevention.

StataCorp. 2001. *Stata Statistical Software: Release 7.0.* College Station, TX: Stata Corporation.

State of California. 1941. *California: Statutes and Amendments to the codes. (Chapter 937).* Sacramento, CA: State of California.

———.1993. *California Youth Authority Committee Report.* Sacramento, CA: State of California.

———. 1996. *California Task Force to Review Juvenile Crime and the Juvenile Justice Response: Final Report*. Sacramento, CA: State of California.

Steffensmeir, D. S., E. A. Allan, M. D. Harer, and C. Streifel. 1989. "Age and the Distribution of Crime." *American Journal of Sociology* 94:803-31.

Steffensmeir, D. S. and M. D. Harer. 1987. "Is the Crime Rate Really Falling? An "Aging" U.S. Population and Its Impact on the Nation's Crime Rate, 1980-1984." *Journal of Research in Crime and Delinquency* 24:23-48.

Sutherland, Edwin H. and Donald R. Cressey. 1978. *Criminology. Tenth Edition.* Philadelphia: J.B. Lippincott.

Taylor, Donald H., Gerda G. Fillenbaum, and Michael E. Ezell. 2002. "The Accuracy of Medicare Claims Data in Identifying Alzheimer's Disease." *Journal of Clinical Epidemiology* in press.

Thornberry, Terence P. 1989. "Reflections on the Advantages and Disadvantages of Theoretical Integration." Pp. 51-60 in *Theoretical Integration in the Study of Deviance and Crime*, edited by S. F. Messner, M. D. Krohn and A. E. Liska. Albany: State University of New York Press.

Thornberry, Terence P. and Marvin D. Krohn. 2001. "The Development of Delinquency an Interactional Perspective." Pp. 289-305 in *Handbook of Youth and Justice*, edited by White. New York: Kluwer Academic/Plenum Press.

Titterington, D. M., A. F. M. Smith, and U. E. Makov. 1985. *Statistical Analysis of Finite Mixture Distributions*. Chichester, UK: Wiley.

Tittle, Charles. 1988. "Two Empirical Regularities (Maybe) in Search of an Explanation: Commentary on the Age-Crime Debate." *Criminology* 26:75-86.

Tittle, Charles R. 1995. *Control Balance Toward a General Theory of Deviance*. Boulder, CO: Westview Press.

Tittle, Charles R. and Harold G. Grasmick. 1998. "Criminal Behavior and Age: A Test of Three Provocative Hypotheses." *The Journal of Criminal Law and Criminology* 88:309-42.

Tolan, Patrick H. and Deborah Gorman-Smith. 1998. "Development of Serious and Violent Offending Careers." Pp. 68-85 in *Serious & Violent Juvenile Offenders*, edited by R. Loeber and D. P. Farrington. Thousand Oaks, CA: Sage Publications.

Tolan, Patrick H. and Peter Thomas. 1995. "The Implications of Age of Onset for Delinquency Risk II: Longitudinal Data." *Journal of Abnormal Child Psychology* 23:157-69.

Tracy, P.E., M.E. Wolfgang, and R.M. Figlio. 1985. *Delinquency in Two Birth Cohorts, Executive Summary*. Washington D.C.: U.S. Government Printing Office.

———. 1990. *Delinquency Careers in Two Birth Cohorts*. New York: Plenum Press.

Tracy, Paul E. and Kimberly Kempf-Leonard. 1996. *Continuity and Discontinuity in Criminal Careers*. New York: Plenum Press.

Vermunt, Jeroen and Jay Magidson. 2000. *Latent GOLD User's Manual*. Belmont, MA: Statistical Innovations, Inc.

Vold, George B., Thomas J. Bernard, and Jeffrey B. Snipes. 1998. *Theoretical Criminology*. New York: Oxford University Press.

Walker, Samuel. 1998. *Sense and Nonsense About Crime and Drugs*. Belmont, CA: Wadsworth Publishing Company.

Wang, Peiming, Iain M. Cockburn, and Martin L. Puterman. 1998. "Analysis of Patent Data -- a Mixed-Poisson-Regression-Model Apprach." *Journal of Business & Economic Statistics* 16:27-41.

Wedel, Michel and Wagner A. Kamakura. 1998. *Market Segmentation*. Boston: Kluwer Academic Publishers.

Wedel, W., W.S. DeSarbo, J.R. Bult, and V. Ramaswamy. 1993. "A Latent Class Poisson Regression Model for Heterogeneous Count Data." *Journal of Applied Econometrics* 8:397-411.

Weitkamp, E.G.M., H.J. Kerner, V. Schindler, and A. Schubert. 1995. "On the "Dangerousness of Chronic/Habitual Offenders": A Re-Analysis of the 1945 Philadelphia Birth Cohort Data." *Studies on Crime and Crime Prevention* 4:159-75.

West, D.J. and Farrington D.P. 1973. *Who Becomes Delinquent?* London: Heinemann.

West, D.J. and D.P. Farrington. 1977. *The Delinquent Way of Life*. London: Heinemann.

Wilson, James Q. and Richard Herrnstein. 1985. *Crime and Human Nature*. New York: Simon and Schuster.

Wilson, Jeffrey, Neal Rojas, Rudy Haapanen, Elaine Duxbury, and Hans Steiner. 2001. "Substance Abuse and Criminal Recidivism: A Prospective Study of Adolescents." *Child Psychiatry and Human Development* 31:297-312.

Winkelmann, Ranier. 1995. "Duration Dependence and Dispersion in Count-Data Models." *Journal of Business & Economic Statistics* 13:467-74.

———. 1997. *Econometric Analysis of Count Data*. New York: Springer.

Wolfgang, M. E., R. M. Figlio, and T. Sellin. 1972. *Delinquency in a Birth Cohort*. Chicago: Chicago University Press.

Youthful Offender Parole Board. 1988. " Response to Reforming the California Youth Authority." In DeMuro et al. (1988) Reforming the California Youth Authority. Bolinas, CA: Common Knowledge Press.

Zimring, Franklin E. 1998. *American Youth Violence*. New York: Oxford University Press.

Zimring, Franklin E. and Gordon Hawkins. 1994. "The Growth of Imprisonment in California." *British Journal of Criminology* 34 (Special Issue):83-96.

———. 1997. *Crime is not the Problem: Lethal Violence in America*. Oxford: Oxford University Press.

Zorn, Christopher J. 1998. "An Analytic and Empirical Examination of Zero-Inflated and Hurdle Poisson Specifications." *Sociological Methods and Research* 26:368-400.