Document Title: Investigating the Simultaneous Effects of Individual, Program and Neighborhood Attributes On Juvenile Recidivism Using GIS and Spatial Data Mining

Author: Philip Harris, Jeremy Mennis, Zoran Obradovic, Alan Izenman, Heidi Grunwald, Brian Lockwood, Joe Jupin, Laura Chisholm

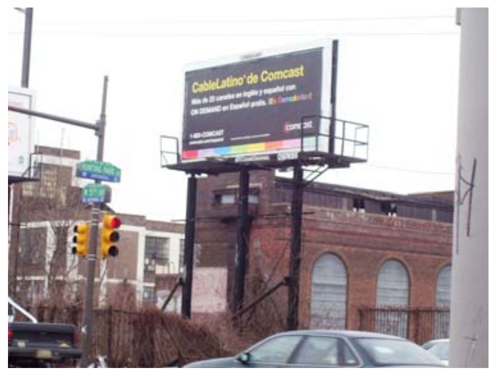
Document No.: 237986

Date Received: March 2012

Award Number: 2006-IJ-CX-0022

INVESTIGATING THE SIMULTANEOUS EFFECTS OF INDIVIDUAL, PROGRAM AND NEIGHBORHOOD ATTRIBUTES ON JUVENILE RECIDIVISM USING GIS AND SPATIAL DATA MINING

Philip Harris, Jeremy Mennis, Zoran Obradovic, Alan Izenman, Heidi Grunwald, Brian Lockwood, Joe Jupin and Laura Chisholm



CONTENTS

THE RESEARCH TEAM

Philip W. Harris is an Associate Professor in the Department of Criminal Justice at Temple University. His research has focused primarily on the areas of juvenile justice, juvenile correctional strategies, and organizational and system development. His current research interests include improving our understanding of how community-based programs interact with neighborhood forces to reduce recidivism and improve other program outcomes among juvenile offenders. Recent publications have appeared in Justice Quarterly, Criminal Justice and Behavior and Evaluation Review. Direct all correspondence to this author at the Department of Criminal Justice, Temple University, 1115 W. Berks Mall, Room 512, Philadelphia, PA 19112, or to phil.harris@temple.edu.

Zoran Obradovic, professor of computer and information sciences and the director of the Center for Information Science and Technology at Temple University in Philadelphia is an internationally recognized leader in data mining. His research focuses on improving predictive modeling and decision support through data-driven discovery and modeling of hidden patterns in large data sets. Obradovic has published about 200 articles addressing data mining challenges in health informatics, the social sciences, environmental management and other domains. His group's pioneering research on the prediction and functional analysis of intrinsically disordered regions in proteins has provided new insight into how protein structure establishes function and the program his team developed was the best rated predictor of intrinsic disorder at three consecutive international competitions organized by protein structure prediction assessment community (CASP 5-7). Obradovic was the program chair at five, track chair at seven and program committee member at about 40 data mining conferences. He currently serves as an editorial board member at seven journals.

Jeremy Mennis is an Associate Professor in the Department of Geography and Urban Studies at Temple University where he specializes in Geographic Information Science. His recent research has focused on modeling human-environment interaction for urban public health and crime applications. Recent publications have appeared in Annals of the Association of American Geographers, Cartography and Geographic Information Science, and Drug and Alcohol Dependence.

Alan J. Izenman is Professor, Department of Statistics, and Director of the Center for Statistical and Information Science, Office of the Vice-President for Research and Strategic Initiatives. His current research interests are primarily in machine learning, data mining, bioinformatics, multivariate analysis (especially high-dimensional data problems), time series analysis, nonparametric function estimation, and the application of novel statistical methods to problems in law and public policy. His graduate textbook entitled Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning, was published in 2008 by Springer. Recent publications have appeared in the Washington University Law Review, Mathematical Reviews, International Statistical Review, Statistical Science, and the Journal of the American Statistical Association. He also wrote three chapters for the edited volume Statistical Science in the Courtroom (J. Gastwirth, ed.).

Heidi E. Grunwald is Assistant Vice Provost for Research and Strategic Initiatives at Temple University. Her current research areas include research design and methodology, hierarchical linear modeling and causal analyses from observational data. Her recent publications have appeared in Research in Higher Education, Journal of Higher Education, and Journal of Behavioral Health Services and Research.

Brian Lockwood is a doctoral student in the Department of Criminal Justice at Temple University. His research interests include the applications of GIS toward the study of crime and near-repeat patterns of offending. A forthcoming book chapter (Rengert and Lockwood, in press) discusses common difficulties that can arise in studies of space and puts forth remedies for those issues.

Joseph Jupin is a doctoral student in the Department of Computer and Information Sciences

Laura Chisolm is a graduate student in the Department of Geography and Urban Studies

ACKNOWLEDGEMENTS

This project was supported by Award No. 2006-IJ-CX-0022 awarded by the National Institute of Justice, Office of Justice Programs, US Department of Justice. The opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Department of Justice.

We wish to thank the following Philadelphia city officials for their support of this project and their assistance in providing us access to data:

Commissioner Anne Marie Ambrose, Department of Human Services, Philadelphia, PA

Commissioner Charles Ramsey, Philadelphia Police Department

Deputy Commissioner Patricia Giorgio-Fox, Philadelphia Police Department

Robert Malvestuto, Chief Probation Officer, Philadelphia Adult Probation and Parole Department

Dr. Ellen Kurtz, Director of Research, Philadelphia Adult Probation and Parole Department

In addition, we wish to thank John Delany, Deputy District Attorney and head of the Trial Division of the Philadelphia District Attorney's Office, and Denise Clayton, Director of the Youth Violence Reduction Partnership, for their support of this project.

ABSTRACT

The primary goal of this project was to develop, apply, and evaluate improved techniques to investigate the simultaneous effects of neighborhood and program forces in preventing juvenile recidivism. For many years, program evaluation researchers have presented the question, "What works to prevent delinquency for whom under what circumstances?" In community settings, answering this question presents a unique challenge, since "circumstances" includes the home neighborhoods of youths participating in correctional programs. Understanding how programs and neighborhoods jointly shape youth behavior and identifying conditions under which rehabilitative programs are successful are fundamental to planning programs that facilitate positive trajectories for physical, social, cognitive, and affective youth development. We investigated the simultaneous effects of neighborhood, program, and individual characteristics (including family) on juvenile recidivism using linear modeling, geographic information systems (GIS) and spatial data mining. GIS provides the technology to integrate diverse spatial data sets, quantify spatial relationships, and visualize the results of spatial analysis. In the context of juvenile recidivism, this approach will facilitate the investigation of how, and why, recidivism rates vary from place to place, through different programs, and among individuals.

The project applies spatial data mining to the analysis of adjudicated juvenile delinquents assigned to court-ordered programs by the Family Court of Philadelphia, Pennsylvania. This population encompasses all adjudicated delinquents committed to programs by the court during the years 1996 to 2002 – more than 26,000 cases. The proposed study makes use of three levels of data: individual, program and neighborhood. In addition to data on individual youths and their families, we will employ a database of designs of the programs that they attended and two or more spatial data sets, including the crime data from Philadelphia Police Department and the U. S. Census. This study includes a vast methodological departure from current practices and can greatly improve the chances of learning more about the dynamics of juvenile recidivism, leading to more effective prevention policies and programs.

CHAPTER 1
INTRODUCTION

Nearly one in four adjudications of youths in juvenile court (23 percent) results in a residential placement (Snyder & Sickmund, 2006). Another 62 percent receive formal probation only; the remaining 15 percent receive a less formal disposition that involves voluntary participation in supervision and related activities. A sizeable portion of probation cases (there are no national estimates) are committed to community-based programs that include mentoring, after-school programming, community service and alternative schools.

Most youths that are placed in residential facilities return to the community and are committed to aftercare programs. In some cases community-based programs deliver services to both youths on an aftercare status and youths committed to them directly. Other youths on an aftercare status are committed to specialized community-based reentry programs. According to national estimates (Snyder & Sickmund, 2006), these youths spend anywhere from two months to one year in placement, with time in placement varying greatly by offense type, gender and race. The 1999 Census of Juveniles in Residential Placement (CJRP) estimated that nearly 100,000 juvenile offenders were released from custody facilities in the United States and returned to their communities (Snyder & Sickmund, 2006). Based on more recent custody counts, it is likely that this number is smaller today, but this number is still likely to add significantly to the demands being place on community-based programs. Compounding the difficulties inherent in the large numbers of juveniles reentering communities is the recidivism rate of these returning juveniles.

Rates of juvenile re-offending have been found to be as high as 66% when measuring recidivism by rearrest and as high as 33% when measuring re-offending by reconvictions within a few years of release (Mears and Travis, 2004; Bureau of Data and Research, 1999). Accurately estimating a national juvenile recidivism is a difficult task, as Snyder and Sickmund (2006: 234) cogently state that, "Such a rate would not have much meaning since juvenile justice systems vary so much across states." The most accurate nation-wide juvenile recidivism statistics are likely to be found by averaging state rates of juvenile recidivism. Further, as evidenced by the varying recidivism rates mentioned above, recidivism rates can differ greatly depending on how recidivism is measured – whether the rate is measuring rearrests, referrals to court, reconvictions, or reconfinement to create the recidivism rate. Using the average of state juvenile recidivism rates for a small number of states, the national juvenile rate could be anywhere between 55% and 25% depending on what measure of recidivism is used to comprise the measure (Snyder & Sickmund, 2006). Regardless of how the concept of juvenile recidivism is operationalized, a sizeable portion of an already large number of juveniles returning from confinement will return to those facilities that once housed them.

A number of studies have identified individual-level predictors of recidivism. These factors include criminal history, age at first arrest, substance abuse, and education (Farrington & Hawkins, 1991; Frederick, 1999; Snyder & Sickmund, 2006; Elliot et al., 1985; Yoshikawa, 1994). Additionally, current age, negative peer relations, family problems, emotional distress, and prior treatment facility placement have been identified as individual-level attributes that increase the risk of juvenile recidivism (Snyder & Sickmund, 2006; Marczyk et al., 2003; Wiebush et al., 2005; Baird, 1984).

In contrast, research examining the effects of community-level variables on rates of juvenile recidivism is less common. Kubrin and Stewart (2006: 167) describe this important, yet overlooked, area of investigation in correctional research by noting that:

Neighborhood context is fundamental to our understanding of why individuals offend, and potentially even more important for understanding why former offenders offend again, yet we know very little about how the ecological characteristics of communities influence the recidivism rates of this population.

They are, however, referring to adult recidivism rather than juvenile recidivism, which has been the focus of even fewer studies of prisoner reentry.

There is good reason to believe that neighborhood attributes have an impact on both juvenile delinquency and recidivism rates. From the pioneering research of Shaw and McKay in the early part of the 20th century to the studies conducted today, space and place have been of particular interest to researchers. A number of studies have found that community context influences several child and adolescent outcomes such as delinquency rates (Curry and Spergel, 1988; Elliott et al., 1996; Herronkohl et al., 2000; Kowaleski-Jones, 2000; Jacob, 2006; Ludwig, et al., 2001; Oberwittler, 2004; Osgood and Chambers, 2000, Sampson and Groves, 1989; Simcha-Fagan and Schwartz, 1986; Wikstrom and Loeber, 2000), IQ (Brooks-Gunn et al., 1993), maltreatment rates (Freisthler, 2006; Freisthler, 2005), health (O'Campo et al., 1997; Sampson, 2001), educational attainment (Brooks-Gunn et al., 1993), and teen pregnancy rates (Brooks-Gunn et al., 2003). Considering these studies and the recent work by Kubrin and Stewart (2006), it is likely that environmental context plays a vital role in determining the rate of juvenile recidivism in a neighborhood.

In the studies reported here, data from a population of juvenile offender cases in Philadelphia were used to investigate four questions: 1) what individual-level variables best predict of juvenile recidivism?, 2) how well do environmental attributes predict juvenile recidivism rates?, 3) do the combined effects of individual and neighborhood attributes improve our ability to predict juvenile recidivism?, and 4) do attributes of the community-based programs delivering services to these youths add any predictive power to individual and neighborhood attributes. The individual level data were taken from the ProDES database,[1] a population database of all juvenile cases committed by the Family Court to community and residential programs between 1994 and 2004. These data were collected to provide program providers, court personnel and funding agencies with trend information and information on program outcomes. From this data set, cases were selected from the years 1996-2002, the years when the data were most complete. In order to test the impact of neighborhood-level attributes on juvenile recidivism, the selected cases were limited to the population of cases committed by the court directly to community programs or committed to an aftercare program following placement in a residential program.

The ProDES data comprise court record data, program intake data, program discharge data, and follow-up data collected six months following program discharge. The program intake and program discharge data were collected by program staff who had been trained by staff of the Crime and Justice Research Center (CJRC) of Temple University, using instruments developed by CJRC. All other data were collected by CJRC staff. Identifying data on juvenile subjects were removed from the database for use by the study researchers.

Data from several other sources, such as the US Census, crime data from the Philadelphia Police Department and a recent health survey, were used in concert with the ProDES data to conduct our analyses. Analyses were conducted to examine the simultaneous effects of individual and neighborhood attributes on juvenile recidivism rates.

We first present the general information about the project, including background literature, data and analytic methods. Because different disciplines and their related analytic methods were applied to the data, we present each of the analytic studies as separate chapters. Chapters 2 through 5 contain the results of studies of the same data set conducted using different analytic methods. Then in a final chapter we present conclusions about juvenile recidivism and methods for studying recidivism, as well as study limitations and implications.

The studies presented in this report were conducted by different sets of researchers who met together frequently to share findings and shared ideas about analytic strategies. The fact that we came from different disciplines was seen as a benefit to the overall research effort. It also meant

[1] ProDES, the Program Development and Evaluation System, was a project of the Crime and Justice Research Center, Temple University, funded by the Department of Human Services, Philadelphia, PA, from 1994-2004. More information on this project can be found at www.temple.edu/prodes .

that choices of which data elements to include in an analysis varied somewhat from study to study. We had not intended this to be the case, but the demands of different analytic tools vary, giving use flexibility in some analyses that did not present itself in others. We make clear these issues and choices at the beginning of each chapter.

Goals of the Study

The impetus for this study was to further explore a set of data on delinquent youths in Philadelphia that had been created initially to support program development and improved matching of youths to programs. This large data set, spanning nearly ten years and comprising 43,000 cases, had never been examined in terms of spatial factors that might be influencing recidivism patterns. This new use of these data implied adding other types of data having to do with the environments in which these youths resided. It also meant selecting time periods in a youth's life when he was exposed to neighborhood environmental forces. Two goals shaped our choices of data and methods:

To investigate the usefulness of geospatial analyses and data mining of social science data to improve knowledge building capacities in juvenile justice

To examine the simultaneous effects of individual, neighborhood and program attributes on recidivism among delinquent youths

Our Research Questions

What benefits are gained from analyzing social science data with geospatial and data mining analytic tools?

How do individual youth and family characteristics, program characteristics, and neighborhood characteristics interact to produce specific program outcomes, such as recidivism and placement in a more secure facility?

Why is recidivism more common in some neighborhoods than others?

Why are certain types of reoffending (offense type) more common in some neighborhoods than others?

To what extent is program impact a function of (constrained and enhanced by) neighborhood forces?

Background to the Study

ProDES (Program Development and Evaluation System) was an outcomes-based information system that tracked the population of Philadelphia's delinquent youths who are court-committed to any type of intervention program. It operated in Philadelphia from January 1994 through December 2004, and was directed by Philip Harris and Peter Jones, both faculty members in the Department of Criminal Justice at Temple University. During its ten years, it accumulated data on over 43,000 juvenile cases. ProDES was designed to provide outcome information to programs for delinquent youths and to users of these programs, namely judges, probation officers and funding agents. Its goals are to provide continual feedback to programs and the juvenile court that will 1) facilitate program development, 2) facilitate better matching of youths to programs, and 3) identify and facilitate improvements in the array of programs available to the Philadelphia juvenile justice system.

ProDES collected data at four points in time: (1) at the point of disposition (the juvenile equivalent of sentencing), data were culled from the youth's Family Court record that contains information such as offense history, placement history, needs (e.g., drug use, mental health problems) and family history; (2) at program intake, staff persons were asked to complete a needs assessment and the youth completes a self-report section containing psychometric scales; (3) at discharge, the intake process was repeated and program staff report on the youth's progress in the program; and (4) six months following program discharge, a follow-up record check was conducted at Family Court by ProDES staff to identify any new petitions (arrests leading to charges) generated in the juvenile or adult court systems, and telephone interviews were conducted with youths, when available, and guardians.

Analyses of these data have included annual program evaluations of more than 100 programs over a period of ten years, creation of a personality typology to investigate how different types of youths respond to specific programs (Harris and Jones, 1999), a study of judicial decisions regarding first-time offenders (Fader et al., 2001), and development of a prediction model to identify youths likely to become chronic offenders (Jones et al., 2001).

Philadelphia neighborhoods, even those characterized by poverty, social isolation, and crime, differ in their ability to protect their youthful residents from making contact with the juvenile justice system. Evidence of this is found in a recent study conducted by Jones et al. (2001), which reported that neighborhoods (as measured by zip code) were the most significant predictor of which first-time offenders predicted to become chronic offenders actually went on to become chronic offenders (accruing three or more arrests). Using the ProDES database, the authors developed a risk instrument identifying the characteristics of chronic offenders at the time of their first contact with the juvenile court.

Fader, et al. (2004), in their study of youths in juvenile aftercare, used the ProDES data and an inventory of community-based youth serving programs supplied by *Philadelphia Safe and Sound* to test two research questions: 1) When controlling for individual-level risk of recidivism, is there significant variation in aggregate rates of recidivism across neighborhoods? 2) Do neighborhoods with the fewest youth-serving programs have the highest rates of recidivism for aftercare clients? Using a binary logistic and complementary log-log (CLL) model, she identified a neighborhood in which aftercare clients were at a very high risk for unsuccessful transition back into the community, and that aftercare programs do not produce the same degree of positive effects for Latino clients as for youths of other racial groups.

Figure 1.1

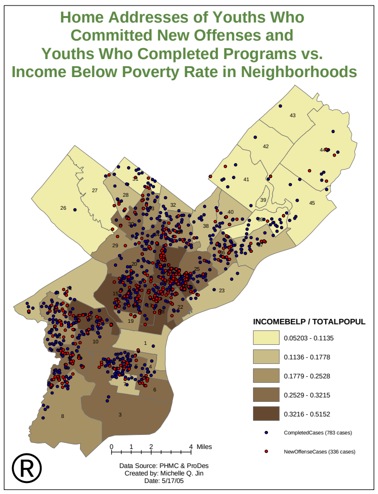


Figure 1.2

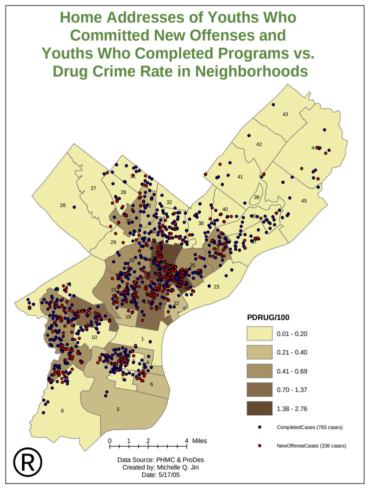


Figure 1.3

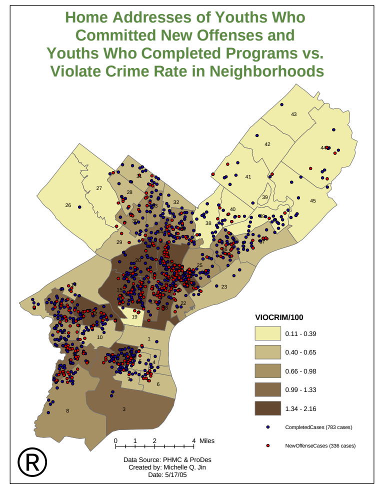


Figure 1 is a preliminary spatial mapping of home neighborhoods of youth recidivism versus income below poverty rates for metropolitan Philadelphia. Figure 2 is a mapping of home neighborhoods of youth recidivism versus drug crime rates and Figure 3 is a mapping of home neighborhoods of youth recidivism versus violent crime rates. Recidivism rates are nested and highest in home neighborhoods with high poverty, drug and violent crime rates. The spatial boundaries in these maps were created through census tracts and local expert opinion.

The Policy Context

Case Processing and Recidivism

This study focuses on youths who have been adjudicated delinquent, who are living at home, and who are participating in community-based programs to which they have been committed by the Family Court. Philadelphia's juvenile justice system and case processing is not substantially different from those of other jurisdictions: police arrest, prosecutors charge, judges adjudicate, probation officers supervise, and correctional programs provide rehabilitation services. For purposes of this study, it is important to emphasize two aspects of the system: 1) judges in Pennsylvania control the decision to commit a youth to a specific program and when to discharge that youth from the program, and 2) Philadelphia possesses a large array of community-based programs ranging from mentoring and after school programs to large alternative schools. The individual-level data set that we used for this study is case based; a case begins with a judge's decision to commit a youth to a program and ends with a judge's decision to discharge the youth from that program. This sequence of decisions can begin with an arrest for an offense, a probation violation, or a change of program due to a case review.

Youths are committed to community-based programs in two ways: 1) a direct commitment that is part of a case disposition following an adjudication of delinquency, violation hearing, or case review, or 2) a step-down from a residential facility, usually referred to as aftercare or reintegration. Thus some youths in community-based programs were committed directly following an institutional stay, while others had never experienced an out-of-home placement.

In the following sections, we describe briefly the components of Philadelphia's juvenile justice system.

Major decision points:

Arrest: the police make an arrest on the basis of evidence

Charging: the District Attorney charges the youth with a crime based on evidence

Detention: the Juvenile Probation Department decides whether to hold a youth in detention pending trial

Pretrial Hearing: A judge reviews the evidence in a hearing to determine whether a trial is warranted or to offer the youth an informal option to a trial

Trial: As in the adult system, this is a fact-finding hearing. However, in addition to a finding of guilt, the judge must find the youth to be a delinquent*

Disposition: A decision by a judge, with input from probation, as to the sanction to be imposed

Case Review: A hearing by a judge in which case progress is reviewed and a decision is made whether or not to change the disposition or to discharge the case from court jurisdiction

Violation Hearing: A hearing by a judge, with input from a youth's probation hearing, pertaining to the youth's failure to comply with the requirements imposed by the judge in the disposition. A violation may result in placement in a more restrictive setting.

*"Delinquent" is a status that includes criminal behavior as well as a judgment that the youth requires supervision above that being provided by the parents. A finding that a youth is a delinquent is usually based on an assessment of the family and the youth's participation in school, as well as prior court involvement. It is possible, although rare, for a youth to be found guilty and not delinquent.

Disposition Options*

*the underlined options are those that pertain to each of our study samples

Probation: All youths will have a probation officer. In a pure probation case, the youth continues to live at home or in the home of a relative, and is supervised by a probation officer (PO). That means that the PO will see the youth, parents, and sometimes teachers, to monitor the youth's behavior.

Foster Care: Placement in an approved foster home for a period of time.

Community-Based Program: The youth living at home and is required by the court to attend a program in the community. The program can be an after-school program, an alternative school, or a mentoring program.

Residential Facility: This and the next three options involve removal of the youth from the home and placement in a facility with other delinquent youths. A residential facility may be a group home in Philadelphia or in the residential area of another city. Typically these youths attend public school.

Non-secure Institution: This is a residential facility that comprises a campus but no measures to prevent the youth from running away other than staff supervision. That is, youths are not locked in their rooms and there are no fences around the facility.

Secure Institution: These facilities are more like minimum security prisons. Youths are locked in their rooms, the buildings they live in are locked, and there is a fence around the facility.

Mental Health Facility (known as a Residential Treatment Facility): These facilities are a mix of secure and non-secure, but they are more heavily staffed and have psychiatrists and psychologists. Most of the youths are on psychotropic medication.

Aftercare: Once a youth completes a required period of time in a residential facility (options 4-7 above), the youth returns to court for a second disposition on the same offense. This is a commitment to an aftercare program, which is a community-based program (like option 3 above) but is specifically designed for youths reentering the community after a period of being incarcerated. We can expect that aftercare cases are more serious cases than those receiving disposition options 1-3.

Literature Review

We are studying the impact of individual, spatial and program attributes on new offenses committed by youths at some point following the disposition decision. Thus our measure of recidivism included only the time that a youth was participating in a community-based program. Presumably, the influence of neighborhood effects was not removed during the period of program participation. In fact, we find that nearly a quarter of our subjects recidivated prior to program discharge. In our analyses, we have made a distinction between:

The youth is charged with one or more offenses that occurred during the time the youth was participating in a program, and

The youth is charged with one or more that occurred during a period of six months following termination of program participation

Our Theoretical Framework

Social Disorganization Theory, macro and micro

Differential Organization Theory

Social Learning Theory

Individual-Level Predictors of Recidivism

As the number of juveniles entering the juvenile justice system increases, so does the need for risk assessment that can inform decision-makers at all stages of the juvenile justice process (Krysik & Lecroy, 2002). As a result, "the use of risk assessment by juvenile justice systems more than doubled between 1990 and 2003" (Schwalbe, Fraser, & Day, 2007). The practice of assessing risk has historically been done via three methods: anamnestic, clinical, and actuarial. Anamnestic methods use historical data to determine the future actions of an individual. Clinical methods involve the human judgment of professionals such as probation officers and psychologists to make risk assessments. Actuarial methods use quantitative analyses of individual characteristics to determine risk. Both clinical and actuarial methods are commonly used today, but studies have shown that actuarial risk prediction consistently outperforms the results of clinical risk prediction (Gottfredson & Moriarity, 2006).

Studies examining the predictors of juvenile recidivism have uncovered a number of individual-level factors that influence the likelihood that a juvenile will re-offend. Research has shown that juveniles at highest risk to offend are those who have done so in the past (Cottle, Lee, & Heilbrun, 2001; Dembo et al., 1998; Farrington & Hawkins, 1991; Frederick, 1999; Snyder & Sickmund, 2006). Other individual-level predictors of recidivism include gender (Dembo et al., 1998; Minor, Hartmann, & Terry, 1997), race (Dembo et al., 1998; Minor, Hartmann, & Terry, 1997), substance abuse (Duncan, Kennedy, & Patrick, 1995; Elliott, Huizinga, & Ageton, 1985; Stoolmiller & Blechman, 2005), early childhood misbehavior (Farrington, 1986; White, Moffitt, Earls, Robins, & Silva, 1990), current age (Snyder & Sickmund, 2006) age at first arrest (Frederick, 1999; Katsiyannis & Archwamety, 1997), education (Dembo et al., 1998; Katsiyannis & Archwamety, 1997; Myner, Santman, Cappelletty, & Perlmutter, 1998; Yoshikawa, 1994) criminal history (Cottle, Lee, & Heilbrun, 2001; Farrington & Hawkins, 1991), prior out-of home placement (Myner, Santman, Cappelletty, & Perlmutter, 1998), peer relations (Akers, 1985; Hoge, Andrews, & Leschied, 1996; Marczyk, Heilbrun, Lander, & DeMatteo, 2003; Myner, Santman, Cappelletty, & Perlmutter, 1998), mental health problems (Huizinga, Loeber, Thornberry, & Cothern, 2000; Pullmann et al., 2006), and family problems (Hoge, Andrews, & Leschied, 1996; Wiebush, Baird, Krisberg, & Onek, 1995). Results of these studies have been used to construct risk assessment tools tasked with assigning levels of risk among juvenile offenders based on the likelihood of re-offending.

The most widely used of these predictive tools is the Youth Level of Service/Case Management Inventory (YLS/CMI). The YLS/CMI, created by Hoge and Andrews (1996) contains eight scales that address criminal history, peer relations, education/employment, family circumstances, substance abuse, leisure/recreation, personality/behavior, and attitudes/orientation. State-wide tools such as the Arizona Juvenile Risk Assessment Form have achieved minimal to moderate success in predicting recidivism (Ashford & LeCroy, 1990), while tools such as the Wisconsin Juvenile Probation and Aftercare Assessment Form have been found to not discriminate between juveniles who do and do not recidivate (Ashford & Lecroy, 1988). Other well-known risk assessment tools include the Psychopathy Checklist-Youth Version (PCL-YV) (Corrado, Vincent, Hart, & Cohen, 2004; Forth, Kosson, & Hare, 2003), the Child and Adolescent Functional Assessment Scale (CAFAS) (Hodges & Kim, 2000; Quist & Matshazi, 2000), and the Model Risk Assessment Instrument (MRAI). Evaluations of these generic risk assessment tools have concluded that these instruments achieve only minimal classification accuracy (Marczyk, Heilbrun, Lander, & DeMatteo, 2003). Similar evaluations have found that these tools do not achieve the same results for juveniles in different jurisdictions and for juveniles of different ethnicity and gender (Miller & Lin, 2007; Schwalbe, Fraser, & Day, 2007; Schwalbe, Fraser, Day, & Cooley, 2006). Notably, these instruments do not include information regarding the neighborhoods in which subjects reside. This is not to say that there is a complete omission of the consideration of community context in the juvenile offender risk literature, because the literature has certainly identified such risks (McCord, Widom, & Crowell, 2001; Shader, 2001), but rather, that many risk prediction instruments do not take geographic or social space into consideration.

Neighborhood/Environmental Predictors of Recidivism

In order to survey current knowledge about the impact of environments on juvenile recidivism, we reviewed the literature in the following areas: tests of social disorganization theory, adult and juvenile prisoner reentry, and risk of recidivism. Since youths in community-based programs are nested in both neighborhoods, each of these areas is instructive as to how environments interact with youth attributes to influence recidivism.

Community-Level Variables and Program Outcomes

Studies of the effects of community characteristics on individual-level outcomes have been conducted for decades. Social disorganization theory is clearly the dominant theoretical framework for these findings. Bursick (1988: 521) defines social disorganization as "…the inability of local residents to solve commonly experienced problems." Shaw and McKay (1971) promoted social disorganization theory through their community-level studies of crime and delinquency in early 1940s Chicago, which borrowed from the earlier work of their colleagues at the University of

Chicago, Park and Burgess. Shaw and McKay (1971) found that income, ethnicity, residential stability, and the physical condition of residences significantly influenced rates of delinquency. Many studies have operationalized the predictors identified by Shaw and McKay using proxies for poverty (percentage of single parent households, percentage living in poverty, percentage on assisted living), residential instability (percentage of residents living in the same home for the last five years and if residents still live within 15 minutes of their childhood home), and ethnic heterogeneity (percentage white, black, Hispanic, other). These studies have found that the major tenets of social disorganization theory to be valid in many different settings today (Sampson and Groves, 1989, Jacob, 2006, Obertwittler, 2004; Osgood and Chambers, 2000; Veysey and Messner, 1999).

It was not until the mid-1990s that there was renewed interest in research on community-level factors and how they affect various outcomes (Sampson et al., 2002). A great deal of the credit for this renewal in community research is attributed to William Julius Wilson's 1987 book, *The Truly Disadvantaged,* that outlined how urban communities have transformed since the 1970s and how these changes have negatively affected these communities. Wilson's work spurred interest in research that examined the role of neighborhood effects in producing a myriad of outcomes, including educational attainment, cognitive skills, and early/unplanned pregnancy (Brooks-Gunn et al., 1993; Elliott et al., 1996; Kowaleski-Jones, 2000; Rankin and Quane, 2000, 2002).

At the same time, a renewed interest in social disorganization theory produced many studies on the role of communities in both increasing and decreasing crime due to informal social control mechanisms and cohesion among neighbors (or a lack of) (Bursick, 1988; Markowitz et al., 2001; Sampson and Groves, 1989; Sampson et al., 1999; Veysey and Messner, 1999). Similarly, the concept "collective efficacy" was developed to capture the extent of community organization, participation, shared beliefs and collective problem solving (Sampson et al., 1999; Sampson et al, 1997). Studies of the effects of collective efficacy have found that rates of crime, delinquency, and other unwanted outcomes are lower in communities with higher levels of collective efficacy.

A number of other measures of community context have been found to be correlated with adolescent outcomes in recent studies. The percentage of families making more than $30,000 annually in a neighborhood (Brooks-Gunn et al., 1993), drug and alcohol availability (Freisthler et al., 2005), violence tolerance of community residents (Oberwittler, 2004), number of unconventional friends a juvenile has (Rankin and Quane, 2002), and organizational participation (Sampson and Groves, 1989; Simcha-Fagan and Schwartz, 1986) have been found to be significantly related to individual outcomes such as child maltreatment, delinquency, and teen pregnancy rates, among others.

Research that measures the effects of community-level predictors on juvenile delinquency are likely to be the most useful when looking at what effects rates of juvenile recidivism. Re-offending by youths already in the juvenile justice system can be expected to be influenced by the very same factors that brought about their initial offending. Our examination of relevant studies, then, includes research on the effects of spatial variables on delinquency. Several of these studies have even specified the type of delinquency that certain neighborhood features are more apt to influence, as research has found that the predictors of different types of delinquency can be different. Jacob (2006) found that residential mobility is the best predictor of juvenile property crime while the rate of lone-parent families is the best predictors of violent crime. The work of Osgood and Chambers (2000) to test social disorganization in rural areas found that the percentage of female-headed households in a neighborhood was the strongest predictor of violent crime committed by juveniles. Sampson and Grove's (1989) test of social disorganization in Great Britain found that their constructs of organizational participation and local friendship groups were the strongest predictors of burglary, while ethnic heterogeneity significantly predicted only property crime (which included vandalism that Sampson and Groves consider to be a crime of juveniles). "Family disruption" was found to predict violent crime and the measure of "unsupervised peer groups" was found to be predictive of both property and violent crime (Sampson and Groves, 1989).

With the recent surge in studies of neighborhood effects on adolescent outcomes has come increasing use of analytical innovations like multilevel modeling, and specifically, hierarchical linear modeling (HLM) (Raudenbush and Bryk, 2002). The between and within-neighborhood effects of neighborhood-level predictors can be examined while controlling for individual-level variables using HLM. This innovative analytical tool and its growing popularity in the studies of the effects of ecological context has provided the means for accurately estimating the role that neighborhoods play in determining the outcomes of adolescents (Elliott et al, 1996; Oberwittler, 2004; Rankin and Quane, 2002).

Prisoner Reentry

The Urban Institute has identified several neighborhood-level variables that characterize communities that former prisoners are likely to return to upon their release. In a study of about 28,000 prisoners returning to Philadelphia neighborhoods after serving time in the Philadelphia Prison system in 2003, a disproportionate amount of former prisoners returned to six of the sixty-nine identified neighborhoods in Philadelphia (Roman et al., 2006). Based on data from the 2000 US Census, these researchers found that these six neighborhoods had several characteristics in common. Three of the communities had lower rates of high school graduates than the city average while five of the neighborhoods had higher rates of residents living below the poverty line than the city average (three of the neighborhoods had poverty rates greater than two times the city average) (Roman et al., 2006). Additionally, all but one of the six neighborhoods had a higher amount of vacant properties than the city average, the average price of residential properties in all neighborhoods was significantly lower than the city average, and three of the six neighborhoods had rates of vacant land parcels that were more than seven times the city average (Roman et al., 2006).

A study of prisoner reentry in Chicago communities found similar trends concerning the existence and characteristics of "target communities" that former inmates were more apt to return to (LaVigne et al., 2003). Of the 15,488 former prisoners that returned to Chicago in 2001, 34% returned to six specific communities of the seventy-seven communities that make up Chicago. Most of the those communities had higher rates of vacant housing, renter-occupied housing, female-headed households, families below the poverty line, Part I crime, and lower rates of high school graduates (LaVigne et al., 2003). The "target communities" that former prisoners returned to at high rates in Chicago, Baltimore, Cleveland, and Houston, where characterized by above average rates of unemployment, female-headed households, and families living below the poverty line. In the cities to which the greatest number of prisoners return to in New Jersey, Virginia, and Massachusetts, the Urban Institute found that these cities have higher than average rates of poverty and female-headed households and poverty rates twice as great as the state averages (Solomon, 2006).

The socially disorganized and resource-deprived characteristics of the neighborhoods that returning prisoners are more likely to return to, as described above, are likely responsible for the high recidivism rates that plague former adult prisoners. In a 2002 study of adult recidivism for prisoners released in 1994, 68% of the 272,111 former prisoners were rearrested for a new crime within 3 years (Langan and Levin, 2002). A three-year follow-up period is an admittedly long period of time, but recidivism rates for studies with one-year follow-ups have consistently discovered recidivism rates of over 30% (Kubrin and Stewart, 2006). Snyder and Sickmund (2006), noting inconsistency in measures of recidivism across states, identified rates of juvenile recidivism on a state-by-state basis with one-year follow-ups. Aggregating data from several states, the juvenile recidivism rate was 55% when defining recidivism as the act of being rearrested, but fell to 33% when requiring that juveniles be reconvicted/readjudicated to be considered recidivists (Snyder and Sickmund, 2006). Pooling together the knowledge gained from the Urban Institute's multi-site study of prisoner reintegration, several neighborhood factors were identified as contributing to the recidivism of the former adult prisoners followed in the research (Solomon et al., 2006). Along with a multitude of individual-level characteristics that were found to increase the likelihood that a returning prisoner would recidivate shortly after his or her release from prison, respondents in the Returning Home study in Illinois who perceived their communities as "safe and good places to live" were at a reduced risk of recidivating and were much more likely to be employed (Solomon, 2006: 15). Additionally, respondents in the Illinois sample who described drug sales as problem in their neighborhood were much more likely to have engaged in substance abuse while in the community, which would seem to predicate higher rates of recidivism than former inmates who did not use drugs after their release (Solomon, 2006).

Juvenile Reentry

The difficulties facing the 100,000 or so juveniles released each year from facilities and the communities to which they return are very similar to those created by the influx of adult prisoners back into communities (Mears and Travis, 2004). Sullivan (2004) describes the phenomenon of the majority of juveniles returning to the same disadvantaged communities from which they came to the facilities from; directly mirroring the trend amongst returning adult former inmates. Lack of adequate aftercare also plagues the success of juvenile reentry, much as it does for adult reentry (Mears and Travis, 2004; Wiebush et al., 2000). In many respects, however, the difficulties for juveniles are more comprehensive due to their young age. For example, while lack of education and work experience are serious barriers for the successful reintegration of former adult inmates back into society, these issues are much more prevalent for teenagers that have not yet had the opportunity to earn a high school diploma or hold a meaningful job (Mears and Travis, 2004).

Psychologically, juveniles are even less prepared for the transitory process of reintegration once they are released from a treatment or correctional facility. Steinberg et al. (2004) describe the "psychosocial maturity" that is necessary in order for a juvenile to transition properly into adulthood that is often lacking or inhibited in juvenile offenders, that compounds the difficulties of successful reentry after a form of juvenile custody. In a similar vein, Snyder (2004) points out that many juvenile offenders suffer from mental disorders, such as schizophrenia and various learning disabilities, that also increase the likelihood of recidivism. Mears and Aron (2003) have estimated that up to 12% of incarcerated juveniles are mentally retarded and slightly more than a third have a learning disability. These mental disorders, due to the young ages of juveniles, are often undiagnosed, and thus, untreated (Mears and Travis, 2004; Mears and Aron, 2003).

Altschuler and Brass (2004) list seven specific domains in which difficulties in juvenile reentry reside: 1) family and living arrangements; 2) peer groups; 3) health; 4) substance abuse; 5) education; 6) employment; and 7) recreation. Several of these areas have already been described, but all are vital to the successful reintegration of a juvenile offender. Consider that more than half of all juvenile offenders have a family member who has served a jail or prison sentence (Snyder, 2004). Or that drug abuse rates, mirroring the adult offender population, are much higher for juvenile offenders than for juvenile non-offenders (Sullivan, 2004). Concerning recreational activities and free time, juvenile offenders are much more likely to spend their time getting into trouble because they have had little practice in using their time constructively (Altschuler and Brass, 2004). This is evidenced by the fact that juvenile offenders released from facilities have spent an average or nearly one-third of their adolescent years incarcerated (Snyder, 2004).

In summary, the well-established ecological tradition in criminology, stemming from the work of Shaw and McKay (1942), has linked crime rates to structural characteristics of communities, including economic, family, and social stability indicators. Subsequent research has found that community context influences many child and adolescent outcomes such as parenting behavior (Chung and Steinberg 2006), maltreatment rates (Freisthler, Gruenewald, Remer, Lery, & Needell, 2007; Freisthler & Merritt, 2006; Freisthler, Needell, & Gruenewald, 2005) problem behavior (Elliott et al., 1996; Rankin & Quane, 2000, 2002), and educational attainment (Ainsworth, 2002; Brooks-Gunn, Duncan, Klebanov, & Sealand, 1993). Other studies have found evidence that health-related outcomes such as low birth weight (O'Campo, Xue, Wang, & Caughy, 1997; Struening, Wallace, & Moore, 1990) child maltreatment (Coulton, Korbin, Su, & Chow, 1995; Freisthler &

Merritt, 2006) and both teenage and non-marital birth rates (Billy & Moore, 1992; Brooks-Gunn, Duncan, Klebanov, & Sealand, 1993) are also influenced by neighborhood-level predictors.

Juvenile crime is similarly related to community context. Many studies have concluded that juvenile crime is dependent on numerous constructs that operationalize neighborhood processes, such as disadvantage and collective efficacy (Beyers, Loeber, Wikstrom, & Stouthamer-Loeber, 2001; Bursik, 1988; Liberman, 2007; Loeber & Wikstrom, 1993; Sampson & Groves, 1989; Sampson, Morenoff, & Earls, 1999; Simcha-Fagan & Schwartz, 1986). Drug and alcohol availability (Freisthler, Needell, & Gruenewald, 2005; Herrenkohl et al., 2000), the spatial concentration of juveniles with delinquent attitudes (Oberwittler, 2004), and the number of "unconventional" friends that a juvenile has (Rankin & Quane, 2002) have also been identified as neighborhood-level predictors positively related to juvenile delinquency. We summarize these studies in Table 1.1.

Table 1.1

Spatial Correlates of Recidivism, Delinquency, and Maltreatment

Factor	**Measure**	**Relationship**	**Source**
Family & Social Networks			
Informal Networks	Proportion of friends living in neighborhood	Informal networks negatively correlated with problem behavior	Elliott et al., 1996.
Neighborhood adults involved in crime	Are neighborhood adults involved in crime?	Higher rates of adults involved in crime lead to more juvenile delinquency.	Herronkohl et al., 2000.
Informal Control	Composed of 4 subscales: mutual respect, institutional controls, social control, and neighborhood bonding	Informal Control is negatively correlated with problem behavior	Elliott et al., 1996.
Organizational participation	Did you participate in a community meeting in the last week?	Significant on burglary rate	Sampson & Groves, 1989
Unsupervised peer groups	Survey items	Significant on all property and violent crime	

Organizational participation	Avg parental education and level of community organizational involvement	Decreases delinquency	Simcha-Fagan & Schwartz 1986
Single-parent families	School-level aggregate of individual-levelfamily structure variable	Schools with higher rates of single-parent families increase delinquency (person crimes)	Anderson, 2002
Local friendship groups	How many of your friends live within 15 mins? (scale of 1-5)	Sig on burglary rate	Sampson & Groves, 1989
Family disruption	Proportion divorced/separated & single-parent households	Sig on violent crime	Sampson & Groves, 1989
Unconventional friends	Scale asking juveniles about how many of their friends are good students and about their attitudes toward school and conventional goals	Level 2 variable of neighborhood disorganization was not found to predict juvenile problem behavior, but unconventional friends did, and UF rates are higher in more disorganized neighborhoods	Rankin & Quane, 2002
Drugs			
Drug possession per population	Rate	Correlated with child maltreatment rates	Freisthler et al., 2006
Drug and alcohol availability	Drug availability constructed using police data on sales and possession arrests. Alcohol availability density created using data from the ABC using number of liquor licenses	Higher levels of drug availability and more bars lead to higher rates of child maltreatment	Freisthler et al., 2005
Alcohol access	Number of places selling alcohol	More bars lead to higher rates of maltreatment	Freisthler et al., 2004

Availability of drugs	Are drugs available in the neighborhood?	Higher drug availiability results in more juvenile violence	Herronkohl et al., 2000.
Poverty			
Various neighborhood-level demographics	Vacant housing, Poverty, HS education, Renter/Owner-occupied housing, median income	Areas that returning prisoners came home to in concentrated numbers have higher levels of poverty and other disadvantage. Similarly, the crime rates in these neighborhoods are higher than average in Philadelphia	Roman et al., 2006.
Affluent families	% families with annual income > $30k	Higher rates of affluent neighbors increases IQ and lowers teen birth rates, and HS drop out rate, which increases negative youth outcomes	Brooks et al., 1993
Neighborhood disadvantage & proportional imbalance of wealth	ND = persons on public assistance, living below poverty line, unemployed, and median family income. Imbalance of wealth uses the ICE index	Both community-level variables are sig related to adult recidivism	Kubrin & Stewart, 2006.
Community-level variables from census	Poverty, owner-occupied home, female-headed households, HS grad, below poverty line, vacant housing, nonwhite	All variables are higher in areas with the greatest concentrations of returning former inmates	LaVigne et al., 2006.
Low-poverty neighborhoods	High = 60% poverty, low = less than 10%	Juveniles moved from high to low-poverty areas committed less juvenile crime	Ludwig et al., 2001

Disadvantaged housing	25% highest scores on a factor composed of SES disadvantage, familism, and stability were designated disadvantaged and lowest 25% as advantaged	Disadvantage did not increase delinquency rates for high risk children, but it did for medium risk children.	Wikstrom & Loeber, 2000
Social disorganization			
Ethnic heterogeneity	Blau index ($1\text{-}\Sigma pi^2$)	Significant on property crime	Sampson & Groves, 1989
Social disorganization/ Impoverishment	% female-headed households, poverty, unemployment, and Hispanics	SD is correlated with higher rates of maltreatment	Freisthler et al., 2004
Social Disorganization & Poverty	Social Disorg. = % of neighborhoods with more than 70% black or white, and neighborhoods with majority Hispanic; Poverty measured with unemployment rate, % living under poverty line, and average mortgage investment	Poverty is strongest predictor of delinquency rate in black neighborhoods, while gang homicide is associated with Hispanic concentration	Curry and Spergel, 1988.
Social disorganization	Census data on SES, mobility, heterogeneity, supervision, and urbanization	SES and instability are best predictors of juvenile crime. Residential mobility is the best predictor of property crime, while lone-parent families is the best predictor of violent crime.	Jacob, 2006.
Residential stability	Proportion living in same home for last 5 years	Residential stability decreases risk-taking behavior and aggressive behavior amongst adolescent, even when controlling for community disadvantage	Kowaleski-Jones, 2000.

Neighborhood quality is also of concern to reentry programs. The majority of returning juvenile offenders re-enter the disadvantaged communities from which they came (Mears & Travis, 2004; Spencer & Jones-Walker, 2004; Sullivan, 2004). Environmental forces such as poverty, housing vacancy, and high residential mobility increase the likelihood of re-offending. Tied in with the return of juveniles to disadvantaged communities is the lack of adequate aftercare that limits the success of juvenile reentry, much as it does for adult reentry (Mears & Travis, 2004; Wiebush, McNulty, & Le, 2000). The effects of these service limitations are likely to be heightened by the quality of the environments in which these youths reside. In that respect, the disadvantaged environments to which juveniles frequently return to are doubly damaging: first, by providing the same environmental context that likely influenced the commission of their initial offense, and secondly, by decreasing the availability of services and programs necessary for successful reintegration.

Risk assessment tools rarely address these environmental forces. (Webster et al., 2006: 12). The research described below will further suggest that risk prediction tools used to assess juveniles would be improved with the inclusion of variables that measure community context.

Program Attributes and Recidivism

A considerable body of research has focus discovering which programs or program types are most effective in preventing recidivism. This work has concentrated on increasing the number of randomized control trials used in meta-analyses to identify types of effective program or program elements in reducing crime and delinquency (Rosenthal and DiMatteo, 2001). Reviews of evaluation studies have also contributed greatly to the growth of knowledge about effective and ineffective program approaches, and the emergence of an "evidence-based program" frame of reference (Greenwood, 2008).

Recent research has concluded that treatment of juvenile offenders can reduce recidivism by as much as 10 percent, which, when considering the tens of thousands of juveniles arrested per year in this nation, can drastically reduce the amount of future crime (Lipsey 1992, 1995). Another meta-analysis of 200 studies of treatment delivered to serious and violent juvenile offenders concluded that treatment reduced recidivism by six percent (Loeber, Farrington, and Waschbusch 1998). As a result, the question has shifted from determining if juvenile offender treatment works to determining what the vital components of effective treatment are and how to replicate them. Research conducted in response to these questions has led to several conclusions that continue to gain support.

One such conclusion pertains to the importance of treatment dosage. The length of stay in treatment programs has been found to be negatively related to recidivism, indicating that as time in program increases, the odds that a juvenile will recidivate decrease (Lipsey 1995, 1999, 1999). Although treatment programs with longer curricula may receive more serious clients at higher risk of recidivism, it is clear that length-of-stay in a program significantly affects the likelihood that a juvenile will re-offend.

Regarding the types of services provided by treatment programs, several studies have identified counseling as a valuable treatment component in the prevention of re-offending among juveniles (Guerra 2008; Lipsey 1999, 1999). The type of counseling identified by these studies as most effective in reducing recidivism is individual counseling. Additionally, programs offering interpersonal skills and skill-oriented services have also been found to reduce recidivism (Lipsey 1992, 1999, 1999).

Family-related services have also been found to be effective components of juvenile offender treatment (Guerra 2008; Lipsey 1999; Gordon and Graves 1995; Henggeler et al. 1998). In particular, multisystemic therapy (MST) is a form of treatment that has been found to significantly reduce recidivism among juveniles by involving the family in the treatment process (Henggeler et al. 1998). MST highlights the importance of including the family of the juvenile offender in the treatment process. Functional family therapy (FFT) is a similar, but less costly, family-related program that has also been found to significantly reduce recidivism among its clients (Gordon and Graves 1995). FFT includes the fostering of problem-solving skills, the increase of emotional bonds between family members, and parental training among its components. Although many current treatment programs do not possess the resources to implement such comprehensive family-related program services, these studies illustrate the importance of involving the family in the treatment process of juvenile offenders.

Multimodal treatment, or programs that offer multiple types of services, have been identified by Lipsey's (1992; Lipsey 1992, 1995, 1999) meta-analyses as programs that garner successful outcomes among their clients. This is an intuitive finding, considering that these programs employ several types of services to rehabilitate their young clients, but it is nonetheless important. Empirical evidence indicates that combinations of intervention methods can achieve more success than programs that employ a single type of service.

Many of the most effective juvenile treatment programs include cognitive-behavioral components. Cognitive-behavioral treatment seeks to modify how offenders think about situations that can lead to offense behavior through the use of one or more intervention components such as anger management, cognitive self-control, social problem solving, social perspective taking, empathy, moral reasoning, and changing attitudes and beliefs (Guerra 2008: 84). The effectiveness of cognitive-behavioral programs has been empirically demonstrated for both adult offenders (MacKenzie 2000; Landenberger and Lipsey 2005) and juveniles (Guerra 2008; Robertson 2000; Izzo and Ross 1990; Robertson, Grimes, and Rogers 2001). In fact, a study of juvenile rehabilitation by Izzo and Ross (1990: 138) concluded "that programs that included a cognitive component were more than twice as effective as programs that did not." Current evidence, then suggests that effective programs for juveniles should include services that attempt to change the way they think.

Few studies have examined how the effectiveness of juvenile treatment components differs by the structure of treatment programs. However, several meta-analyses of juvenile treatment programs have determined that the level of success and effective components of juvenile treatment differ between institutional and noninstitutional programs (Lipsey 1999, 1999; Lipsey and Wilson 1998). Lipsey and Wilson (1998), in their meta-analysis of 200 programs for serious and violent juvenile offenders, concluded that community-based programs garnered larger effects than did institutional programs. For example, individual counseling and behavioral treatment components have been found to be much more effective for noninstitutionalized youth than for juveniles who reside in a treatment facility (Lipsey 1999; Lipsey and Wilson 1998). This suggests that individual counseling and behavioral treatment is more likely to be effective in community-based treatment programs than in residential institutions. Further support for the success of individual counseling in community-based programs is the finding that individual counseling is more effective in community-based programs than is group counseling (Guerra 2008).

Studies that have examined the program-level correlates of treatment for specific populations of juveniles have largely come to the same conclusion: cognitive-behavioral treatment is a component that is frequently associated with success. This is true for serious and violent offenders (Altschuler 1998; Lipsey and Wilson 1998), sex offenders (O'Reilly and Dowling 2008) and offenders suffering from co-occurring issues of mental health and substance abuse (Veysey 2008). Regardless of the type of juvenile client, juvenile treatment programs should have some type of service or services that address the way juvenile offenders think about their offending and their world.

The value of evaluation studies rests on the knowledge that program designs have been implemented properly. The literature on program implementation is clear: most programs fail not because of bad design but because of faulty implementation (Ellickson and Petersilia 1983; Harris and Smith 1996; Mihalic et al. 2004). Meta-analyses of juvenile offender treatment by Lipsey (1995; Lipsey 1999) have similarly identified the importance of implementation in program success, concluding that the size of the researcher's role in an intervention is positively correlated with successful outcomes. Van Voorhis and colleagues have identified a number of deficiencies in programs that they and others have reviewed that take us well beyond a program's treatment components (1997: 276-279). Included in their list of deficiencies are poorly qualified staff, ignorance of research evidence regarding treatment methods, and a lack of appropriate clinical and risk assessments. These deficiencies underscore the lack of capacity of many programs to design and implement effective programs.

Programs are, of course, more than services delivered to clients. When we look closely at programs, we find complex systems of organizations, organizational environments, treatment approaches, staff characteristics, goals and objectives, intervention approaches, service components and delivery systems, and facilities. As Palmer (1994: 161) has argued, to label a program in terms of any one characteristic is likely to inhibit our ability to use the program effectively and to understand its performance. Program outcomes, however, are not simply a product of a set of intervention activities. This view implies that a set of activities delivered anywhere by any appropriately qualified people to any eligible clients should produce the same results. In reality, programs differ in terms of organization, location, size, staffing, management style, target population, physical resources, and geographical area served.. These “non-intervention” components are critical to a program’s capacity to be successful and will be examined in this research.

Knowledge regarding the effectiveness of intervention programs for delinquent youths is still in its infancy. Research findings have moved from the “nothing works” perspective of the 1970’s, associated most often with Martinson (1974) to a “something works” position that has received strong support from Palmer (1992), Gendreau and Ross (1987), Andrews et al. (1990), Lipsey (1992) and others. Both literature reviews and meta-analyses point to patterns in impact evaluations that indicate strengths and weaknesses of different modalities, but, as Cullen and Gendreau (2000) have noted, what we have at present is enough evidence to support further research. Scholars have yet to answer the question, “What works to prevent delinquency for whom under what circumstances?”

Previous Research with ProDES

ProDES was designed to provide outcome information to programs for delinquent youths and to users of these programs, namely judges, probation officers and funding agents. *ProDES* (Program Development and Evaluation System) is an outcomes-based information system that tracks the population of Philadelphia’s delinquent youths who are court-committed to any type of intervention program. Its goals are to provide continual feedback to programs and the juvenile court that will 1) facilitate program development, 2) facilitate better matching of youths to programs, and 3) identify and facilitate improvements in the array of programs available to the Philadelphia juvenile justice system. *ProDES* collects data at four points in time: (1) at the point of disposition (the juvenile equivalent of sentencing), data are culled from the juvenile’s record that contains information such as offense history, placement history, needs (e.g., drug use, mental health problems) and family history; (2) at program intake, staff persons are asked to complete a needs assessment and the youth completes a self-report section containing psychometric scales; (3) at discharge, the intake process is repeated and program staff report on the youth’s progress in the program; and (4) six months following program discharge, a follow-up record check is conducted to identify any new petitions (arrests leading to charges) generated in the juvenile or adult court systems, and telephone interviews are conducted with youths, when available, and guardians. Analyses of these data have included annual program evaluations of more than 100 programs over a period of ten years, creation of a personality typology to investigate how different types of youths respond to specific programs (Harris & Jones, 1999), a study of judicial decisions regarding first-time offenders (Fader et al., 2001), and development of a prediction model to identify youths likely to become chronic offenders (Jones et al., 2001). Not all of this research is relevant to the proposed study but the latter one clearly is.

Philadelphia neighborhoods, even those characterized by poverty, social isolation, and crime, differ in their ability to protect their youthful residents from making contact with the juvenile justice system. Evidence of this is found in a recent study conducted by Jones et al. (2001), which reported that neighborhoods (as measured by zip code) were the most significant predictor of which first-time offenders predicted to become chronic offenders actually went on to become chronic offenders (accruing three or more arrests). Using the ProDES database, the authors developed a risk instrument identifying the characteristics of chronic offenders at the time of their first contact with the juvenile court. Of the youths who had chronic-prone characteristics, none of the youths living in the 19144 zip code (Germantown) went on to become chronic offenders. On the other hand, more youths in the 19143 (Kingsessing) zip code became chronic offenders than were predicted. The authors concluded that these extreme differences were likely a result of differential access to neighborhood resources.

Fader (2004), in her study of youths in juvenile aftercare, used the ProDES data and an inventory of community-based youth serving programs supplied by Philadelphia Safe and Sound to test two hypotheses: 1)When controlling for individual-level risk of recidivism, there will be significant variation in aggregate rates of recidivism across neighborhoods; 2) Neighborhoods with the fewest youth-serving programs will have the highest rates of recidivism for aftercare clients. While she found no systematic neighborhood effects, she did find that three findings that are significant for local human service and juvenile justice policymakers. First, it has identified a neighborhood (North Philadelphia East-19133) in which aftercare clients are at a very high risk for unsuccessful transition back into the community. A second, possibly related finding is that aftercare programs are somehow not producing the same degree of positive effect for Latino clients as for youths of other racial groups. It is possible that the very people responsible for carrying this work out – the front line social work and probation staff – have some important contributions for developing a more coherent theory about this distressing phenomenon. This type of question seems especially appropriate for examination through ethnographic methods, especially since there is a dearth of literature that allows for youths and families involved in the juvenile justice system to discuss their experiences (for important exceptions, see Fox and Benson, 2000; Hil and McMahon, 2001).

The Data

Data for this study derive from several sources, and comprise three nested levels of data: juveniles, within programs, within residences in Philadelphia neighborhoods. The juvenile data (*ProDES*) are case-based, with a 7-year sample of 26,464 cases (10,980 juveniles) with cases in family court between 1996 and 2003. The juvenile sample is comprised of three qualitatively different groups: juveniles committed to out-of-community, residential programs, juveniles sent to in-community, non-residential programs, and those participating in mandatory aftercare programs after release by residential programs back into the community. Although the sample ranges in age from 10 to 20 years old, the majority (69%) are between 15 and 17 years old. These cases are primarily male (90%) and African-American (73%).

The individual level data were taken from the ProDES database,[2] a population database of all juvenile cases committed by the Philadelphia Family Court to community and residential programs between 1994 and 2004. These data were collected to provide program providers, court personnel, and funding agencies with trend information and information on program outcomes. The ProDES data comprise court record data, program intake data, program discharge data, and follow-up data collected six months following program discharge. The data include measures of family demographics, juvenile characteristics, criminal history, current offense characteristics, recidivism status, and many other items. The program intake and program discharge data were collected by program staff who have been trained by staff of the Crime and Justice Research Center (CJRC), using instruments developed by CJRC. All other data were collected by CJRC staff. Identifying data on juvenile subjects were removed from the database for use by the study researchers.

The program level data consist of various data elements on 109 juvenile treatment programs that received funds from the Philadelphia Department of Human Services. There are several types of programs, including aftercare, in-community school programs, drug and alcohol treatment, counseling, boot camps, institutions, and state detention centers. Data was collected on the structure and location of the program, its target population, its staff, rejection and removal criteria, objectives, and service delivery.

The neighborhood level data consists of two parts. The Philadelphia Health Management Corporation's 2002 Household Health Survey provides data on residents' perceptions factors indicating collective efficacy and neighborhood functioning. 2002 census data allows for the use of information including racial and socio-economic status of neighborhoods, residential stability indicators, and crime rates. A forty-five neighborhood shapefile is used, with boundaries determined by the Philadelphia Health Management Corporation's work with focus groups and planning commissions. The offenders are very much clustered geographically, with 10% of the total from the Hunting Park, 6% from Paschall/Kingsessing, and about 5% each from Strawberry Mansion, Mill Creek, Nicetown/Tioga, Olney, and Overbrook. It is clear that crime rates and indicators of disadvantage are not scattered randomly throughout the city. About 24% of the residents in each neighborhood had incomes below 150% of the poverty line; however, poverty is clustered, with at least 40% of residents in areas such as Hunting Park, Poplar/Temple, Mill Creek,

[2] ProDES, the Program Development and Evaluation System, was a project of the Crime and Justice Research Center, Temple University, funded by the Department of Human Services from 1994-2004. More information on this project can be found at www.temple.edu/prodes .

Upper Kensington, and W. Kensington falling below the poverty line. Neighborhood race is also clustered; residents in areas such as Strawberry Mansion, Mill Creek, and West Oak Lane were over 95% non-white, while residents in some Northeast neighborhoods (Fox Chase, Mayfair, Holmesburg, and Bridesburg) were less than 10% non-white. High rates of drug crimes and violent crimes are clustered in Upper and West Kensington, Poplar/Temple, Nicetown/Tioga, and Sharwood/Stanton.

The Philadelphia Police Department also provided data for this study. These data included type and location of all crime in the city of Philadelphia from 2000-2002, excluding rape, and contain 321,785 crime events occurring during that two-year period. The data were divided into eight crime types: homicide, robbery, assault, burglary, theft, vehicle theft, weapon violation, and drug crime. ArcView GIS was used to geocode these crime events onto the map of Philadelphia neighborhoods. Of the 321,785 crime events in the police data, 299,855 were successfully geocoded for a success rate of more than 93% - well above the 85% minimum success rate for geocoding crime data set forth by Ratcliffe (2004).

Individual-Level Predictors

Data Selection

We began with a total of 26,464 individual delinquency cases. These were cases of youths adjudicated for a delinquent offense and committed to either a residential or community-based program. A case did not require a new adjudication; each new disposition, including a change in commitment due to program completion or probation violation, constituted a new case. Population characteristics are summarized in Table 1.2.

Missing Data

In the vast majority of data sets collected by government, industry, or for academic research purposes, we expect some data values to be missing. For example, consumer package goods data, which include scanner data, tend to have 5% to 15% missing values each week.

In our case study, there are several different types of missing data. Data were collected chronologically in this study and arranged in seven parts, each part of which was obtained over a period of six months to two years, depending upon the case. The seven types of data are: (1) court record data, (2) program staff assessment, (3) initial self-reporting assessment, (4) discharge staff assessment, (5) discharge self-reporting assessment, (6) court-recorded recidivism, and (7) telephone interview six months following discharge.

Table 1.2

	Males	Females	Unknown	Total
Ethnic Category				
Hispanic/Latino	3379	288	0	3,667
Non-Hispanic	20,396	2,401	0	22,797
Unknown	0	0	0	0
Total	23,775	2,689	0	26,464
Racial Category				
American Indian/Alaskan	0	0	0	0
Asian	364	16	0	380
Hawaiian/Pacific Islander	0	0	0	0
Black or AA	17,277	2,001	0	19,278
White	2,654	378	0	3,032
More than 1	0	0	0	0
Unknown/Not reported	3,480	394	0	3,874
Totals	23,775	2,789	0	26,464
Hispanic-Racial Category				
American Indian/Alaskan	0	0	0	0
Asian	0	0	0	0
Hawaiian/Pacific Islander	0	0	0	0
Black/AA	0	0	0	0
White	0	0	0	0
More than 1	0	0	0	0
Unknown	3,379	288	0	2,667
Totals	3,379	288	0	2,667

In most of these data sets, we find missing data. There are different reasons for cell values to be missing. In (1), certain data were never recorded. In (2), (3), (4), and (5) some of the assessment data were not received from the program because of a simple lack of cooperation from program management, or staff turnover (42% turnover per year on average), or because the youth stayed with the program fewer than 30 days. In (7), the youth's phone number may not have been correct, or no answer was obtained following multiple attempts to contact the youth, or because the youth refused to be interviewed. There are no missing data in (6).

When data are missing, the software must reflect those facts, treating those spaces appropriately. Most relational databases incorporate the NULL as a means of indicating the absence of a data value, which might mean that the value is missing, unknown, nonexistent (no observation could be made for that entry), or that no value has yet been assigned. A NULL is not equivalent to a zero value or to a text string filled with spaces. Sometimes, zeroes replace missing values, other times by estimates of what they should be, based upon the rest of the data. In our case study, missing data are denoted by a blank or 9, 99, 999, or 9999, depending upon the point of entry and what would be a feasible answer (e.g., for a question on number of children, the missing symbol cannot be 9; so it is recorded as 99). If items were missing from the forms when the data were entered, the coders recorded the missing data as 9 or 99 or 999. Such inconsistency in coding has to be dealt with prior to the analysis stage.

The effect of missing data on a statistical analysis depends upon how many missing values there happen to be, where they occur, whether there is a discernable pattern to their missingness, and in which variables (fixed or free, input or output) they appear. How we treat such missing data should certainly play a major role in any analysis and should have an impact upon any inferences derived from the data.

For the data sets selected for this project, we found that missing data varied with data category. In particular, clinical information such as self report scales and needs assessment information was missing in approximately 45 percent of cases. We were, therefore, forced to exclude these variables in spite of their theoretical usefulness. In other studies, it may be preferable to sacrifice external validity in favor of data inclusion. Given the size of the sample, it would be possible to estimate selection bias if such a decision were made.

Of the total sample, we found the following with regard to missing data:

- ProDES Data
    - Court Record data (prior court case records, demographics): Less than 1% missing
    - Self-report scales at Program Intake (self-esteem, values, family bonding, school bonding): 35 to 45% missing
    - Staff-reported scales at Program Intake (needs assessment): 35 to 45%
    - Discharge from Program: 13% of cases are missing all discharge data
    - Of cases with some Program Discharge data:
    - Staff reported scales at Discharge: 44% missing
    - Self-reported scales at Discharge: 66% missing
- Program data: Less than1% missing data

- Neighborhood data: No missing data

Imputing Missing Data

The art of dealing with missing observations in the analysis of a data set had never assumed the significance that it does today. A number of simplified methods have been used to circumvent the nuisance of missing data values, including one popular method which deletes those observations that contain missing data and analyzes only those cases that are observed in their entirety (often called complete-case analysis). We may regard this simple strategy as similar to what might be done if we discovered outliers in the data. Such a complete-case analysis may prove to be satisfactory if the proportion of deleted observations is small relative to the size of the entire data set and if the mechanism that leads to the missing data is independent of the variables in question --- an assumption referred to by Rubin as "missing at random" (MAR) or "missing completely at random" (MCAR) depending upon the exact nature of the missing-data mechanism (Little and Rubin, 1987). Any deleted observations may be used to help justify the MCAR assumption. Of course, this assumes that the domain expert can identify the missing-data mechanism.

In some situations, either the missing data constitute a sizeable proportion of the entire data set or we cannot justify the missing data as MCAR. If there are extensive missing data, then it is possible that very few cases (or even none at all) may remain after deleting those cases that contain missing values. Single-imputation-based procedures have been used to impute (or "fill in") an estimated value for each missing observation, and then analyze the amended data as if there had been no missing values in the first place. Sometimes, the singly-imputed value is just a mean of all the completely-recorded values for that variable (mean imputation); at other times, it is a value predicted by a regression on the completely-recorded data (regression imputation). The biggest drawback of using a single-imputation method such as these is that sampling variability due to imputing the missing value cannot be incorporated into the analysis as an additional source of variation, which, therefore, leads to underestimating the standard errors of model estimates.

Since the late-1970's, Rubin and his colleagues have introduced a number of sophisticated algorithmic methods --- such as the EM algorithm (Dempster, Laird, and Rubin, 1977; Little and Rubin, 1987) and multiple imputation (Rubin, 1987, 1996) --- for dealing with incomplete data situations, especially for large public-use data sets from sample surveys and censuses. These methods have since become commonplace in applied statistics research and practice.

Multiple imputation (MI) is a data-based statistical method for dealing with situations in which there are a modest amount of missing data. MI takes each missing datum and imputes it using several values, each of which represents an acceptable substitute value for the missing datum. A single imputation creates a revised data set, which can then be analyzed as if the imputed values were the real values obtained from those cases where the missing values occur. This imputation procedure is repeated a small number m of times, where m is usually between 5 and 10. These imputations create m different data sets and, hence, m different model estimates. These m model estimates are then averaged and a measure of standard error of this averaged model estimate is derived. These estimated standard errors, thus, reflect the additional variability absorbed by the model estimates when the imputed values are incorporated into the data analysis.

Rubin's MI technique is essentially a Bayesian method: A parametric model is specified for the complete data as well as a model that specifies the mechanism by which data become missing; next, a prior distribution for the unknown model parameters is specified; then we simulate M independent draws from the posterior distribution (i.e., the conditional distribution of the missing data given the observed data) by Bayes's Theorem. To ease the computational burden in nontrivial applications, special tools, such as Markov chain Monte Carlo (MCMC) procedures are used (Schafer, 1997).

Even if one carries out MI using very simplistic methods, it will produce results that are far superior to any other equally-easy method to implement (e.g., complete cases, single imputation) because the multiple copies of the data set allow the uncertainty about the missing data to be incorporated into the final inferences (Heitjan and Rubin, 1990).

Rubin has shown that, under certain conditions, MI leads to "frequency-valid" results. Studies have indicated that MI also tends to be quite robust when truth departs from the imputation model (e.g., when dealing with binary or ordered categorical variables).

In our case study, missing data will be filled in by using a multiple imputation procedure; the specific details depend upon the type of statistical model we intend to apply to the data. In the statistical modeling and data mining scenarios of our case study, the spatial relationships of the data will be of special importance. These data are then considered as spatial data and will be analyzed as such. As a result, the missing data will have to be imputed by taking into consideration their spatial characteristics.

It is generally recognized that the problem of imputing missing values in spatial data is a particularly difficult one. The fact that values may be missing has special significance in a data set where interest focuses on the arrangement properties of the data and where a missing value refers to a particular area or location, rather than as a missing value from an experiment which contains replicated data.

Some attention has been paid to techniques for imputing missing values in spatial data. For example, Griffith, Bennett, and Haining (1989) used maximum likelihood to impute missing cells in spatial data, with an application to urban census data, and Dass and Nair (2003) used Bayesian hierarchical models and spatial smoothing to impute missing multivariate spatial data for image reconstruction.

Format of missing data in our data set was either system-missing (blank), 9, 99, 999 or 9999 depending on point of entry and feasible answers (ex. For number of youths, missing cannot be 9 so is 99). If items were missing from forms when entered, coders entered 9/99/999. If the entire form was missing, the value was usually system missing, but some variables were recoded in final cleaned database.

A New Cluster-Based Method for Imputing Missing Values

The ProDES data contains a large amount of missing values. Imputing these values by standard methods is a problem since the data is a mixture of categorical and continuous attributes. To address this problem and related problems in other social science data we proposed a dynamic clustering imputation (DCI) algorithm relies on similarity information from shared neighbors, where mixed type variables are considered together. When evaluated on a public social science dataset of 46,043 mixed type instances with up to 33% missing values, DCI resulted in more than 20% improved imputation accuracy over Multiple Imputation, Predictive Mean Matching, Linear and Multilevel Regression, and Mean Mode Replacement methods. Data imputed by 6 methods were used for test of NB-Tree, Random Subset Selection and Neural Network-based classification models. In our experiments classification accuracy obtained using DCI-preprocessed data was a lot better than when relying on alternative imputation methods for data preprocessing. A manuscript with detailed experiments of this study is included as an attachment (under review at 11th International Conference on Data Warehousing and Knowledge Discovery).

Data Reduction

The individual level data were taken from the ProDES database, a population database of all juvenile cases committed by the Philadelphia Family Court to community and residential programs between 1994 and 2004. These data were collected to provide program providers, court personnel, and funding agencies with trend information and information on program outcomes. The ProDES data comprise court record data, program intake data, program discharge data, and follow-up data collected six months following program discharge. The data include measures of family demographics, juvenile characteristics, criminal history, current offense characteristics, recidivism status, and many other items. The program intake and program discharge data were collected by program staff who were trained by staff of the Crime and Justice Research Center (CJRC), using instruments developed by CJRC. All other data were collected by CJRC staff. Identifying data on juvenile subjects were removed from the database for use by the study researchers.

In order to test the impact of neighborhood-level attributes on juvenile recidivism, the selected cases were limited to the population of cases committed directly to community programs or committed to an aftercare program following placement in a residential program. Our aim was to include only cases for which neighborhood forces had a potential direct effect on recidivism. Based on that criterion, 13,000 cases were selected from the period between 1996 and 2002 - the years when the data were most complete. The data set was further reduced by the removal of females, as prior research has demonstrated a gender difference concerning the predictors of juvenile delinquency and recidivism (Daigle, Cullen, & Wright, 2007; Funk, 1999; Mazerolle, 1998). These considerations resulted in our first sample for analysis which consisted of an all-male, juvenile population that had been committed to programs within their communities by the Philadelphia

Family Court (n=11,036). Excluding youths in foster homes and group homes reduced this number to 10,971. These 10,971 cases included multiple records for the same youth; representing each time he received a new program committment. Approximately one third of the sample was composed of repeat offenders with the majority appearing in the dataset twice, and approximately 300 appearing more than twice. Because there were not enough observations within each youth to allow for a longitudinal analysis, we decided to select the first occurring case for each youth resulting in a sample size of 7,282 male juveniles.[3] Of these boys, 2,565, or 36 percent, were on aftercare status, meaning that they had been returned to the community after spending time in a residential setting. This difference between youths who were on aftercare and those committed to community programs directly was accounted for by including a measure of aftercare status in the analysis.

Some variation in numbers of cases will be observed across analytic methods. These differences are due to two main factors: requirements of the method regarding the handling of missing data, and, in the case of the HLM analysis with program data, exclusion of cases in programs with numbers of cases less than five. The analysis involving geographical data has a lower number of cases that the number used in the HLM analysis (7166 vs. 7282) because of 116 youths that were spatial outliers. The following table shows numbers of cases and percent missing cases for each of the analyses. The Plaid and Neural Networks analyses used a slightly smaller data set (n=6675) because these analyses involved much larger numbers of variables. Cases with more than 30 percent missing data were excluded.

Table 1.3 Numbers of Cases and Missing Data

All Data Files	#Fields	Missing	%Missing
Level 1 youths and programs (n=7061)	797893	39543	4.9559%
Level 1 youths (n=7282)	888404	23443	2.6388%
Sensitivity Analysis (n=11,036)	408332	1201	0.2941%
Plaid and Neural Networks (n=6675)	2963700	174056	5.8729%
Spatial analysis cases (n=7166)	766762	2	0.0003%

[3] Although we selected the earliest community program experience juvenile in our dataset, this does not necessarily indicate that the case selected represents their first petition to the Philadelphia Family Court. Rather, it represents the first chronological instance in which they were committed to a community or aftercare program by the Philadelphia Family Court during the study period of 1996-2002.

Neighborhood/Environmental Predictors

Data Selection

The original neighborhood level data included 2000 Census data, crime location data from the Philadelphia Police from 2000-2002, and aggregated neighborhood resident perception data from the 2002 Philadelphia Health Management Corporation (PHMC). The ProDES system includes home addresses for each juvenile corresponding to where they listed their home addresses after completing the community or residential treatment program ordered by the Family Court. ArcView GIS 9.2 was used to geocode the home addresses of the juveniles.

We first utilized the neighborhood boundaries delineated by the Philadelphia Health Management Corporation (PHMC), which exhaustively partitions the city into 45 neighborhood polygons. The PHMC surveys biannually a sample of Philadelphia residents within each of the 45 neighborhoods and includes items related to neighborhood safety and perceptions. Preliminary hierarchical linear models using the PHMC neighborhoods, however, indicated that none of the neighborhood perception variables were significant predictors of youth recidivism. We believe that the level of aggregation was too large, resulting in the capture of several distinct populations within each neighborhood, thereby washing out the effects within the neighborhoods and reducing the variance between neighborhoods.

We considered approaches that would allow us to aggregate neighborhoods so that the spatial units would be small enough to capture the immediate environmental influences around a youth's residence, while at the same time were large enough to contain enough cases so that reliable estimates of the intra-neighborhood variance could be calculated in our analysis. The use of Census tracts as proxies for neighborhoods has been an accepted practice in studies of neighborhood context (Elliott et al., 1996; Kubrin, Squires, & Stewart, 2007; Kubrin & Stewart, 2006). The city of Philadelphia is divided into 381 Census tracts, which make for substantially smaller spatial units when compared to the 45 neighborhoods of the PHMC. An initial analysis of the number of cases per Census tract (n=318) indicated that the data was too sparse when nested in Census tracts. Of the 381 Census tracts, 83 contained zero cases, while an additional 114 tracts contained fewer than 20 cases each. Adjacent tracts were merged based on similarities in race and socioeconomic status, and taking into account major barriers such as rivers, railroads, and major highways, so that each spatial unit contained a minimum number of cases sufficient for supporting the reliable estimation of intra-neighborhood variance. The resultant neighborhood file contained 210 neighborhoods. Figure 1 illustrates the new neighborhood boundaries overlaid onto the 381 Census tracts of Philadelphia.

The dotted lines indicate boundaries of the Census tracts from which the new neighborhoods were derived. Dotted lines visible within the solid-lined, grey-filled neighborhoods illustrate neighborhoods that have been created by merging two or more Census tracts. The mean of juvenile addresses across the 210 neighborhoods is 34.6. Thirty-eight (38) neighborhoods had counts of juvenile addresses below 20, with 11 juveniles being the fewest count per neighborhood. These neighborhoods were included in the study, and were not merged to increase the number of cases, in order to preserve theoretically interesting community attributes. Demographic data from the 2000 Census and Philadelphia Police Department were subsequently aggregated to the 210 neighborhoods.

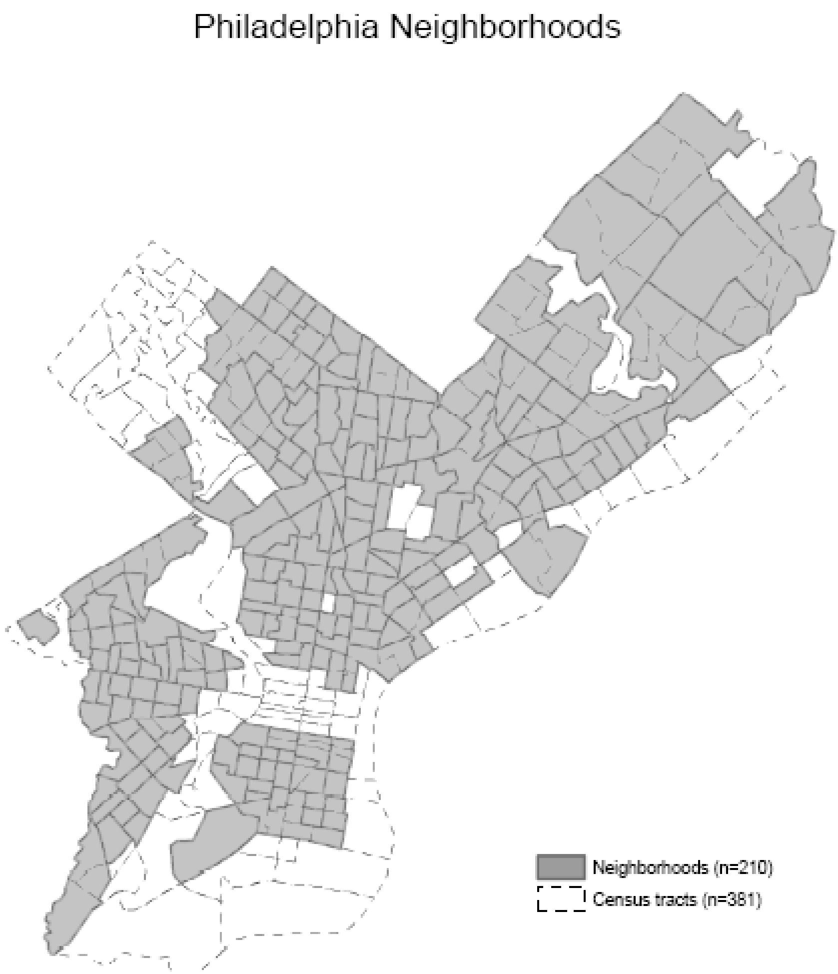
The Philadelphia Police Department provided data on the type and location of all Part I crime events in the city of Philadelphia from 2000-2002, excluding rape. This data set contains 321,785 crime events occurring during that two-year period. The data were divided into eight crime types: homicide, robbery, assault, burglary, theft, vehicle theft, weapon violations, and drug crime. ArcView GIS 9.2 was used to geocode these crime events onto the map of the 210 Philadelphia neighborhoods. Of the 321,785 crime events in the police data, 299,855 were successfully geocoded for a success rate of more than 93%, a figure that is greater than the 85 percent minimum success rate for geocoding crime data put forth by Ratcliffe (2004).

Data Reduction

Initially, subject addresses were geocoded onto a map of 45 Philadelphia neighborhoods delineated by the Pennsylvania Health Management Corporation (PHMC). The PHMC conducts research on health-related outcomes for their bi-annual Southeastern Pennsylvania Household Survey (SPHS) and has aggregated Philadelphia into 45 distinct neighborhoods to understand possible contextual effects on the outcomes in question (http://www.phmc.org/chdb/householdsurvey.html, 2007). This level of spatial aggregation was desirable not only because of the careful and deliberate consideration by the PHMC in the construction of the 45 neighborhoods of Philadelphia, but also because it would allow the current study to utilize the data gathered by the SPHS at the neighborhood-level. However, analyses using the 45 neighborhood-level of aggregation indicated that the neighborhoods might be too large and result in the capture of several distinct populations within each neighborhood, thereby reducing meaningful variance between those groups.

The study then looked at Census tracts as the geographical-level of aggregation. The city of Philadelphia is divided into 381 Census tracts, which make for substantially smaller spatial units when compared to the 45 neighborhoods of the PHMC. Additionally, using Census tracts as proxies for neighborhoods has been an accepted practice in studies of neighborhood context (Kubrin et al., 2007; Kubrin & Stewart, 2006; Elliot et al., 1996). An analysis of the number of cases (n=10,971) per Census tract (n=318) indicated, however, that the data is too sparse when nested in Census tracts. Of the 381 Census tracts, 83 contained zero cases, while an additional 114 tracts contained less than 20 cases each. With a threshold of 20 addresses per tract in order to maintain statistical significance during analysis, many of the Census tracts were removed from the analysis, while several tracts were merged to form larger areas containing more than 20 juvenile addresses. To complete the process, the 69 neighborhoods of the Philadelphia Neighborhood Information System, an initiative of The University of Pennsylvania's Cartographic Modeling Lab, were used as a guide in the merging of Census tracts with low juvenile address counts to approach the threshold of at least 20 cases per tract (http://cml.upenn.edu/nis/index.html). The resultant neighborhood file contains 210 neighborhoods (n=210). Figure 1 illustrates these new neighborhoods overlaid on the 381 Census tracts from which the neighborhoods originated.

Figure 1.4



The dotted lines indicate boundaries of the Census tracts from which the new neighborhoods were derived. Dotted lines visible within the solid-lined, grey-filled neighborhoods illustrate neighborhoods that have been created by merging two or more Census tracts. The mean of juvenile addresses per neighborhood is 52.2. Three neighborhoods have counts of juvenile addresses below 20, with counts of 19, 19, and 14, but these neighborhoods were permitted to be included in the study due to theoretical concerns that these neighborhoods be allowed to exist as they are and not be merged with other neighborhoods. Demographic data from the 2000 Census were retrieved and aggregated up to the 210 neighborhood-level used in this study.

Program Attributes

From 1999 to 2004, the Crime and Justice Research Center (CJRC) collected information every one to two years on the designs of intervention programs serving Philadelphia adjudicated youth. The Program Design Inventory (PDI) collects and stores program-level information such as: mission, objectives, target population, services, activities, and staffing characteristics. Preparing for data collection entailed an extensive literature review of the key components of criminal and delinquency programs. CJRC staff looked at studies involving process evaluations, program design and implementation. In particular, the Correctional Program Assessment Inventory (Gendrau and Andrews, 1994) provided a framework for describing the programs. Incrementally they added key variables until a complete list was formulated.

Data collection involved two components – the first part was a survey completed by a program administrator; the second part involved a face-to-face interview with one or more program administrators. The survey included questions on funding, licensing, and staff characteristics (including gender, race, age, education, training hours, languages spoken and turnover).

The interview between program administrator(s) and CJRC staff included items covering

- contact information
- organizational structure
- facility (if relevant) and size of program
- program goals and objectives
- program activities (including dosage and location of delivery)
- target population (including demographics, geographic area, needs, and offense types)
- reasons for rejection or removal from the program

Data Selection

To estimate effects of components of community-based programs on juvenile recidivism, the selected cases were limited to the population of juveniles committed to community-based treatment programs or to aftercare programs following placement in residential programs. Cases were selected from the period between 1996 and 2002 when the data were most complete.

Females were removed from the eligible population due to findings of gender differences in the predictors of both juvenile delinquency and recidivism (Daigle, Cullen, and Wright 2007; Funk 1999; Mazerolle 1998). Based on these criteria, the resulting population consisted of 10,971 cases. An examination of these cases determined that approximately one third of the juveniles appeared in this population appeared more than once, with the majority appearing twice. As the number of observations per youth would not allow for a longitudinal study, we elected to select the first case per juvenile from within this population of cases. The resulting population contains of 7,282 male juvenile offenders.

Analytic Strategies

Hot spot analysis was useful for offering a visual presentation of recidivism patterns in the data. Since we were primarily concerned with prediction, we viewed this preliminary step as one that could demonstrate the spatial nature of re-offending and indicate relationships to be explored with other methods. In this section, we describe briefly those methods used to improve our understanding of recidivism.

Logistic Regression

For multivariate analyses in which we examined the potential spatial dependency of type of re-offense, we employed stepwise-forward logistic regression to test the association of the explanatory variables with each of our four outcome variables. This approach is a well-established method for reducing the number of explanatory variables in a regression model by iteratively adding explanatory variables to the regression equation only when their relationship with the outcome is significant, after taking into account the influence of the other explanatory variables already present in the model (Darlington, 1990). Such an approach aids in the development of parsimonious models and interpretation of the regression. The stepwise procedure was carried out using four blocks of explanatory variables, where block 1 consisted of the variables describing the characteristics of the individual juvenile, block 2 entered the juvenile's instant offense type, block 3 consisted of variables indicating the social disorganization of the juvenile's residential neighborhood, and block 4 consisted of the spatial contagion variables. If an explanatory variable is entered into a model during the stepwise procedure in one block, it is kept in the models for the subsequent blocks.

Hierarchical Linear Modeling

HLM has become a valuable analytic tool in criminology as researchers have begun to examine social disorganization theoretical concepts at the individual level. The work of Robert Sampson best exemplifies this trend (see, e.g., Sampson, Raudenbush, & Earls, F.,1997). Other researchers interested in social organization's impact on individuals have used HLM to examine violent crime and social organization (Browning, Feinberg, & Deitz, 2004), childhood violence among African Americans (Stewart, Simons, & Conger, 2002), adolescent development (Cook, Herman, Phillips, & Settersten, 2002), and birthweight (Morenoff, 2003). At the same time, we have very few program evaluation studies that have made use of this analytical method (Osgood & Smith, 1995; Duncan & Raudenbush, 1999). Instead, the primary debate in evaluations of programs for delinquent youths centers on the value of experimental methods.

Geographically Weighted Regression

GWR can be considered an exploratory statistical technique that allows a researcher to investigate the nature of spatial-nonstationarity. GWR builds on the expansion method (Casetti, 1972, 1997) to account for more complex local variation in model parameter estimates. Whereas conventional regression generates a single equation to represent global relationships among variables, GWR calibrates the regression equation differently for each observation based on a unique weighting of all observations. Consider the conventional regression equation

$$y_i = a_0 + \sum_k a_k x_{ik} + \varepsilon_i$$

Geographically weighted regression modifies this equation so that there is an individual parameter estimate for each observation's location. The equation may thus be rewritten as

$$y_i = a_0(u_i, v_i) + \sum_k a_k(u_i, v_i) x_{ik} + \varepsilon_i$$

where (u_i, v_i) represents the coordinate location of observation i (Brunsdon, Fotheringham, and Charlton 1996; Fotheringham, Brunsdon, and Charlton 1998; Fotheringham, Brunsdon, and Charlton 2002). Calibration of GWR takes place by weighting all observations according to a distance decay function away from observation i. Let w_{ij} stand for the weight of observation j for the GWR calibration centered on observation i. A Gaussian function may be used to calculate w_{ij} as a continuous function of distance such that

$$y_i = a_0(u_i, v_i) + \sum_k a_k(u_i, v_i) x_{ik} + \varepsilon_i$$

Co-Cluster Analysis

The "plaid" clustering algorithm was introduced in Lazzeroni and Owen (2002). In that paper, it was applied to a food nutrition data set (961 different foods, 6 nutritional components), a foreign exchange data set (18 currencies of different countries over 276 months), and a gene expression data (taken from 10 experimental series involving 2467 genes). All these data examples had continuous variables. Some follow-up work has started to appear on plaid models; see Turner, Bailey, Krzanowski, and Hemingway (2005), who apply plaid models to gene expression data and repeated-measures data, and Calda and Kaski (2008), who provide a Bayesian version of plaid. See also Izenman (2008, Section 12.8.2).

For each example in Lazzeroni and Owen (2002), the data array was partitioned into “layers” of two-way ANOVA models. Each “layer” is formed from a subset of the rows and columns and can be viewed as a two-way clustering of the elements of the data array, except that rows (individuals) and columns (variables) can be members of different layers or of none of them. Hence, overlapping layers are allowed. For example, in the gene expression (microarray) data, the plaid algorithm searched for layers in which clusters of genes would match up best with the different experimental series. For the food nutrition example, 10 layers were found; for the foreign exchange example, 3 layers were found; and for the gene expression example, 40 layers were found.

This idea of searching for overlapping layers consisting of subsets of variables matched with subsets of individuals is unknown in social science research. Furthermore, to our knowledge, the plaid algorithm has not been applied to binary-valued data. Fortunately, the plaid algorithm makes no explicit assumption that the variables have to be continuous or have to be generated from a particular probability distribution. The plaid program was downloaded from the website http://www-stat.stanford.edu/~owen/clickwrap/plaid, where a manual for the software can also be found.

The plaid model can be written approximately as a sum of several terms: an overall effect term for the entire data set together with a sum of terms called “layers.” Each term in the sum consists of a weight function times the product of two indicator functions. One of the indicator functions is equal to 1 if the ith variable is in the kth layer, and is zero otherwise. The other indictor function is equal to 1 if the jth individual is in the kth layer, and is zero otherwise. So, a term will only be present in the sum if both indicator functions equal 1; that is, if both the ith variable and the jth individual are simultaneously in the kth layer. The weight function of each term in the sum can be expressed in a variety of different ways, but here we use the two-way additive ANOVA representation of a layer effect plus a row effect plus a column effect. To avoid overparametrization, a specific requirement to the model is added so that the sum of the row effects for each layer equals zero and the sum of the column effects for each layer equals zero.

A criterion Q is used to estimate the various unknown plaid model parameters from the data. This criterion is an error sum-of-squares criterion, where each term is the squared error in using the plaid model to predict the observed entry in a particular row and column, summed over all r columns and all n rows. Given a number K of layers, the optimization problem quickly becomes computationally infeasible: each row or column can be in or out of each layer, which means that there are (2^r-1)(2^n-1) possible combinations of rows and columns to consider. To resolve this computational problem, the minimization of the criterion Q is turned into an iterative process, where we add one layer at a time. We omit the technical details, which can be found in Izenman (2008, Section 12.8.3).

Data Mining: Neural Network Analysis

Noise, low signal and missing data are major limiting factors that adversely affect the analysis of social sciences data. This is especially true with data that is aimed at recording human behavior, which is quite random. However, it is also understood that the environment affects human behavior. The data used for this analysis are a mix of case data about the child and the child's family, the treatment program that the child attended, and spatially generated environmental data about the child's neighborhood. We have investigated various data mining techniques to increase the accuracy of both predictions (neural networks are tested with unseen cases) and modeling (neural networks are tested with training data). We hope to reduce noise and/or increase signal by partitioning the data into smaller, more focused sets. The target variables for the analysis are drug recidivism (XDrugs), person recidivism (XPerson), property recidivism (XProperty) and all recidivism (ganypet).

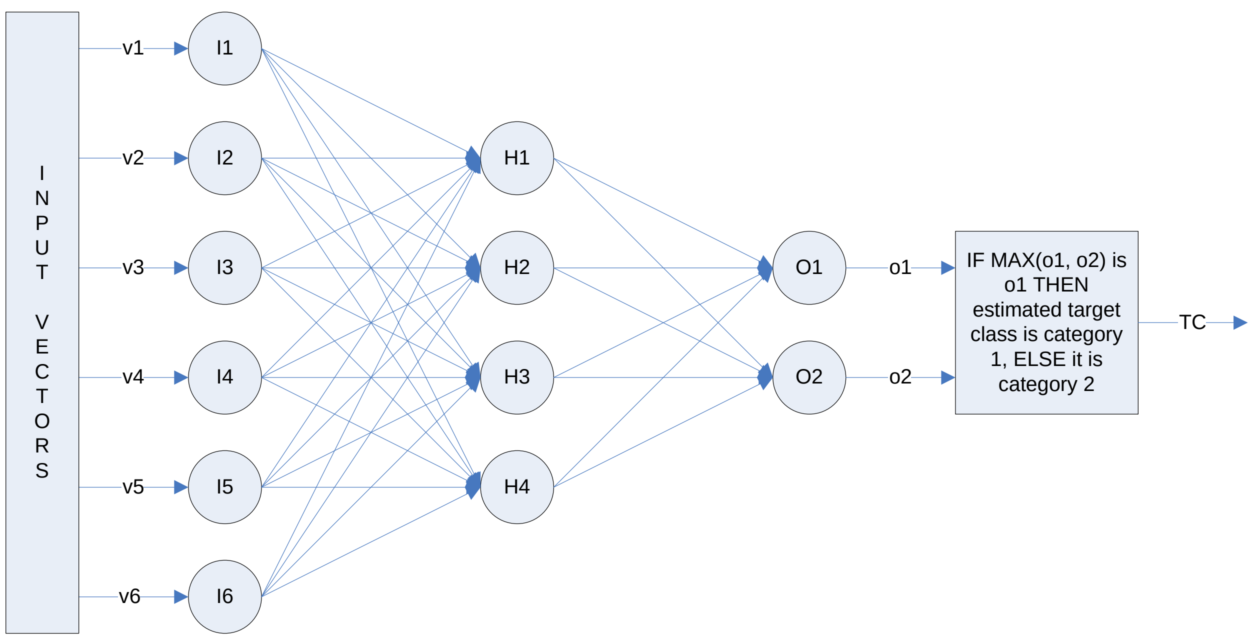
Neural networks are one of the most powerful tools currently used in data mining. They are extremely effective universal nonlinear function approximators, which are able to model functions that are too complex for regression or decision trees. The abstract structure of a typical neural network is that of a fully connected digraph. The most common type, which is used in this experiment, is a feedforward network. They contain multiple layers of nodes that are connected by weighted links. This arrangement was inspired by the neurons and synapses that are found in the human brain. Feedforward networks only allow data to flow in one direction, forward, through the network. Attributes are applied to the input layer, are processed within the network, and the results are rendered from an output layer.

Generally, these networks have an additional set of bias values associated with each processing layer of nodes (input nodes generally don't transform data). These values are handled as an additional attribute weight with constant input value, usually 1. They help to balance the operation by their independence from the input values. Each processing node sums the values of all incoming weights and usually contains a squashing function. This is typically a sigmoid function, hyperbolic tangent or a threshold function. The sigmoid function scales the summed weights between "0 and 1", the hyperbolic tangent between "−1 and 1", and the threshold function produces either "0 or 1" or " −1 or 1", depending on the whether the summed values are greater than, less than or equal to the threshold value. This function output is the processed output of the node.

The back propagation algorithm is used in training these networks. A set of training instances is repeatedly fed through the network through possibly hundreds or thousands of iterations, calculating an output and processing the errors via an error function. This function adjusts the weighted connections towards a value that produces a desirable output. A typical error function takes the derivative of the sigmoid function, solves for 0 to find minima for the instances, preferably global, and adjusts the weight and bias to calculate this derived value. This is called a gradient descent algorithm because it should reduce the distance between the network's estimate and the actual target value, similar to skiing down a mountain. The idea is that the effect is gradual and should improve with each iteration until it converges on an optimal set of estimates for the inputs target classes and terminates. Error functions have momentum and learning rate values to control the adverse effects of erratic instances. The learning rate controls the amount of change that can occur in any given correction. The momentum preserves the overall direction of adjustment, giving consideration to past corrections.

Multilayer perceptrons (MLPs) are a class of back propagation feedforward neural network that possess additional hidden layers, not accessible from outside the network. In its initial state, before any processing occurs, random values are assigned to the weighted connections. An instance set of numeric attributes is fed into the network's input layer of neurons. The values are transformed by the weighted connections and produce an output that can model complex functions because of the additional power provided by the hidden layers. The error function adjusts the weights until some terminating condition is encountered. The graphic, Figure 1.5, depicts an abstraction of an MLP. It shows a network with four input variables in the input layer, a hidden layer with four hidden nodes and an output layer with two output nodes. The data fed into this network would have four variables and three target classes.

Figure 1.5.



Summary

These analytic methods mark a unique contribution of this project to the study of juvenile recidivism. One of our primary goals was to determine if knowledge development can be enhanced by analyzing the same set of data with methods derived from different disciplines. We chose methods that were specifically useful for analyzing spatial data; in some cases the method had been rarely applied to social science data.

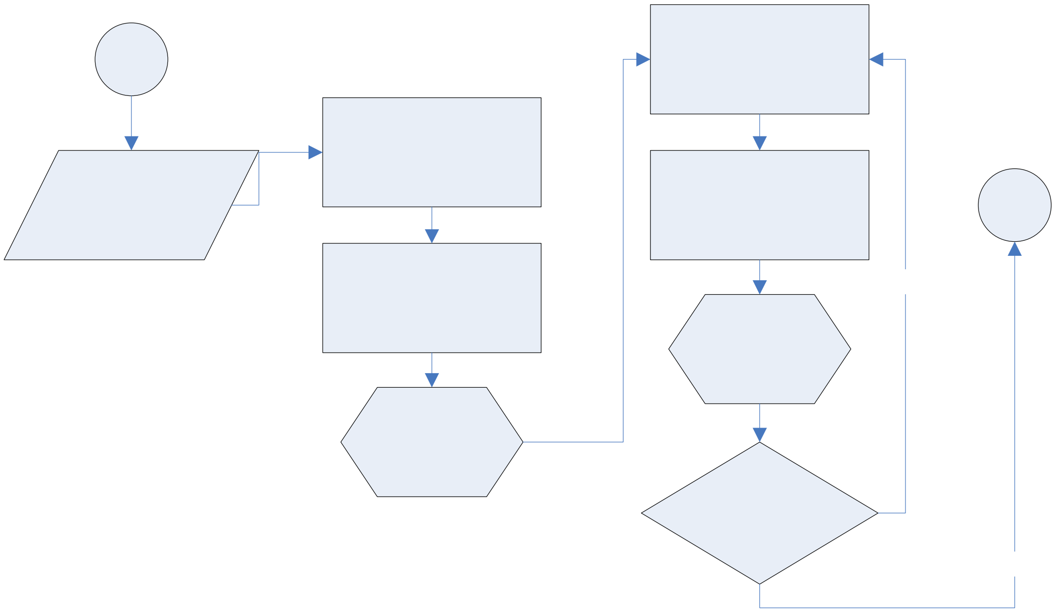
In the following chapters, we examine the ProDES and environmental data using a variety of analytic methods, some of which are unfamiliar to social scientists and others that will be new even to scholars in computer sciences. Our aim is to explore the potential contributions of these methods and to develop hypotheses about juvenile recidivism that can serve as a foundation for subsequent research.

Sensitivity Analysis using Neural Networks with Environmental Data

We began this study with the assumption that spatial data would add value to efforts to predict juvenile recidivism. It may be that individual-level attributes are sufficient to explain recidivism and environmental characteristics add very little to the power of predictive models. To test this assumption, we employed a neural network analysis.

Sensitivity analysis was performed using neural networks to find significant variables in data sets that include a subset of ProDES, aggregate values based on this subset and adult crime data from the Philadelphia Police Department. Sensitivity analysis is performed by training a neural network on a set of data and observing the models accuracy when each predictor variable column is replaced with a column of normally distributed random numbers. The most significant variables for the trained model will produce the greatest decrease in accuracy.

Figure 1.6



Neural networks are very effective at exploiting multidimensional relationships between predictor variables. However, one problem with using neural networks is that they are non-determinant. Two neural network models trained from the same data will typically produce different classifications. This is due to the initial randomization of weighted connections within the network before training, which can lead to undesirable models that have low classification accuracy. In most cases the network will converge towards an optimal model. In some cases it may not. To decrease or even eliminate the effect of undesirable models, we used committee machines, which are composite networks of multiple trained models that poll each model for an estimate of a target value for each case. Figure 1.7 below shows a committee machine with four inputs and five neural networks.

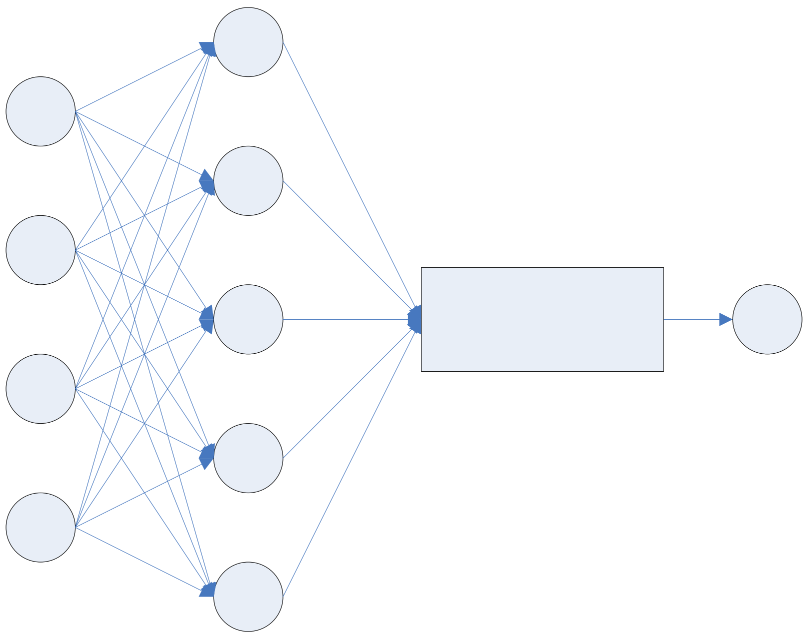
Start

Ne
C

Data Set
With n
Variables
And 1 Target

E

Figure 1.7



The figure below shows a comparison of the accuracies of the average accuracies of five trained neural networks and the accuracies of the committee machines for each randomized column. The average accuracy of the committee machines is 8 percent higher than the averages of five machines. The two rightmost set of bars are the average of all the random tests and the baseline.

NN1

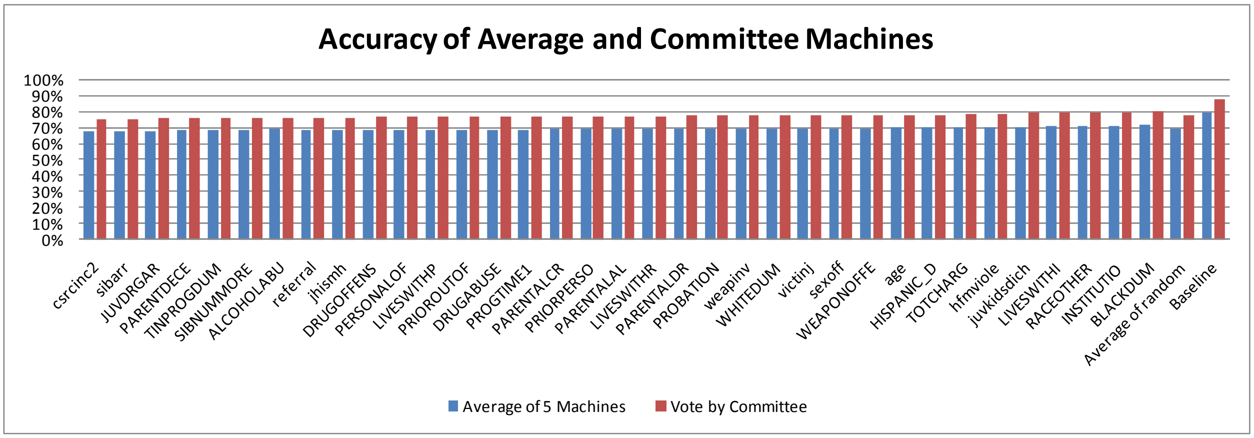
I1

NN2

I2

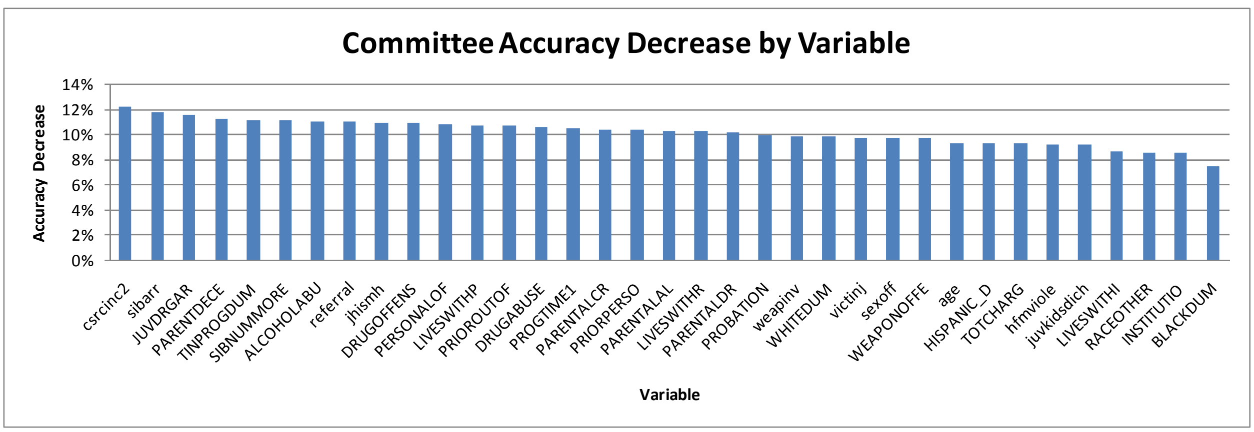
NN3

Figure 1.8



The figure below shows the differences in accuracies for the initial subset of 35 ProDES variables used in this experiment. It is sorted by attributes that had the most affect on accuracy. For the baseline test, it is clearly evident that there is some difference in the decrease of accuracy due to the randomization of these variables.

Figure 1.9



The variable definitions are:

Family had public assistance income at instant offence (Csrcinc2)
Any siblings arrested (Sibarr)
History of juvenile drug arrests (Juvdrgar)
Is at least one of the juvenile's parents deceased (Parendece)
Dichotomized time in program (Tinprogdum)
Juvenile has more than two siblings (Sibnummore)

Juvenile history of drug alcohol (Alcoholabu)
Any record of DHS referrals (Referal)
Does the juvenile have a mental health history (jhismh)
Instant offense was drug related (Drugoffens)
Instant offense was person related (Personalof)
Juvenile lived with parent at instant offense (Liveswithp)
Has the juvenile had out of home placements (Prioroutofhomepl)
Juvenile history of drug abuse (Drugabuse)
Time in program (Progtime1)
At least one parent has a criminal history (Parentalcr)
History of juvenile arrest for person related offences (Priorperso)
At least one parent has had an alcohol abuse history (Parentalal)
Juvenile lived with relatives other than parent at instant offense (Liveswithr)
At least one parent has had a drug abuse history (Parentaldr)
Juvenile on probation at instant offence (Probation)
A weapon was involved at instant offense (Weaponinv)
Juvenile is white (Whitedum)
Victin injured at instant offense (victinj)
Instant offense was sex related (Sexoff)
Instant offense was weapon related (Weaponoffe)
Age of juvenile (age)
Juvenile is Hispanic (Hispanic_d)
Total number of charges (Totalcharge)
History of family violence (Hfmviole)
Juvenile has children (Juvkidsdich)
Living arrangement: institution (Liveswithi)
Juvenile is not black, white, Asian or Hispanic (Raceother)
Juvenile lived in institution at instant offense (Institution)
Juvenile is black (Blackdum)

In addition to the variables above, aggregate variables were calculated for each listed above in various sized concentric rings around the child's residence. The ring sizes are: 33, 66, 100, 200, 300 and 400 meters. The sums of all 35 variable's values were calculated for all youths within each radius of each ring, centered on each child's residence at instant offense but excluding the current child's data. This was also done for the target value representing recidivism, measured as any new petition to Family Court between the disposition decision and six months following discharge from the program to which the youth was committed (ganypet). Similar aggregate values were also calculated for the adult crime data and child crime data based on crime type. These were calculated for:

Drug

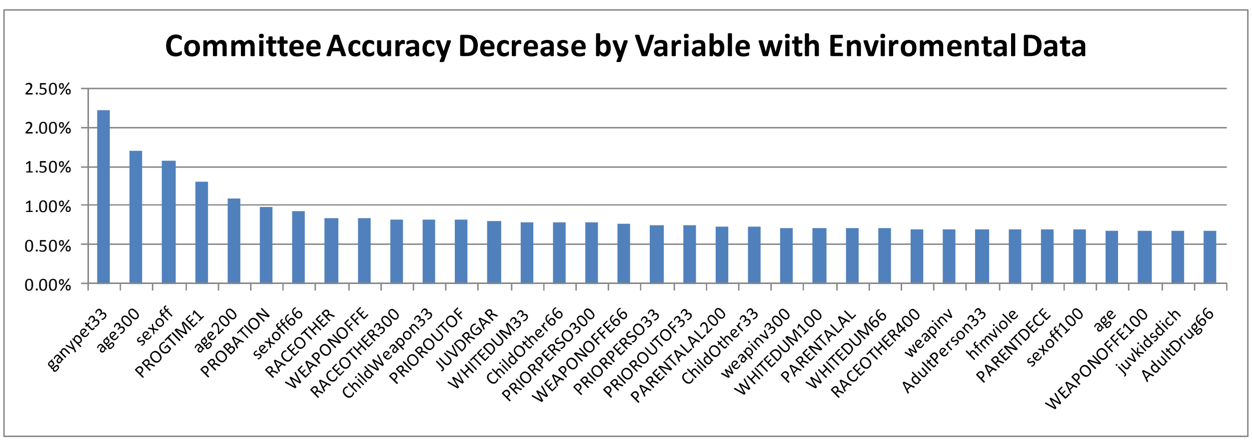
Person
Property
Weapon
Other crimes

Aggregate counts were obtained for the following **adult** crimes:

Drug
Person
Property
Weapon

The figure below shows the results for the data set containing the initial 35 variables, the six sets of aggregates based on the original 35 variables and the target (216 variables), six sets for the juvenile crime types (30 variables) and six sets for the adult data (24 variables). The aggregate variables are represented as variableNameX, where the ProDES variable name has a number X, representing environmental aggregation at X meters as described above.

Figure 1.10



Note that 22 of the 35 most significant variables are environmental aggregate variables. The most significant variable is number of juvenile recidivists within 33 meters of the current child's residence (Ganypet33). Age of juvenile offenders within 200 and 300 meters are both in the five most significant variables. The distribution of this experiment is very different from the baseline (original 35 variables above). This is evident by the sharp decrease in the first seven variables, after which it seems to almost level off. The regression model below only contains the seven top variables from the neural network committee analysis. The model shows that living near other juveniles that are more likely to recidivate and being on probation at instant offense increase the likelihood of recidivism, and that sex offenders are less likely to recidivate.

Table 1.4

0.0845	ganypet33
0.0725	PROBATION
0.0056	age200
0.0021	PROGTIME1
-0.0058	age300
-0.06	sexoff66
-0.172	sexoff
0.3458	Intercept

The C4.5 decision tree algorithm was applied to the data using the Weka data mining software package. The tree had 867 nodes with 434 leaves, which is not really general enough to be used for a decision tool. C4.5 uses a measure called information entropy to find the most significant variable to split at each level in the tree. Entropy uses expected information to partition the data set at each split. Consider a set of training examples S with each sample being a tuple of variables and a target class, where there are m target classes and s_i samples for target C_i for $i = 1, \ldots, m$. The probability that an arbitrary sample belongs to class C_i is $\frac{s_i}{s}$, where s is the total number of samples in the set. Expected information is defined as:

$$I(s_1, s_2, \ldots, s_m) = -\sum_{i=1}^{m} \frac{s_i}{s} \log_2 \frac{s_i}{s}$$

Entropy uses expected information on a set of variables X with values $\{x_1, x_2, \ldots, x_v\}$ to create a partition on S, which contains subsets $\{S_1, S_2, \ldots, S_v\}$ with S_j, containing the samples from S that have value x_j from X. Entropy is defined as:

$$E(X) = \sum_{j=1}^{v} \frac{s_{1j} + \cdots + s_{mj}}{s} I(s_{1j}, \ldots, s_{mj})$$

At the root node of the C4.5 tree, the variable with the greatest entropy is used to make the first split. Subsequent splits continue to further partition the data.

HotSpot, an association rule mining algorithm, was used to analyze the data. HotSpot discovers tree-like structured rules for a given target class that focus on only one target category, maximizing the potential accuracy for rules based on the category of interest. This algorithm allows the selection of rules to be generated based on a minimum support (number of instances considered for the rule) and the increase in accuracy gained by adding a new rule. The rules below are for modeling non-recidivism. The first set requires that a rule consider a support level of at least 16.6% (>=1832) of instances and that the addition of another term in the rule increases accuracy by at least 1%. The second set of rules for non-recidivism lower the accuracy gain to ½%.

Table 1.5

```
Node 1

========

Total population: 11036 instances

Target attribute: ganypet

Target value: 0 [value count in total population: 6432 instances (58.28%)]

Minimum value count for segments: 1832 instances (16.6% of total population)

Maximum branching factor: 2

Minimum improvement in target: 1%


ganypet=0 (58.28% [6432/11036])

 PROGTIME1 <= 12 (62.92% [2026/3220])

 |  PRIOROUTOF33 <= 0.3333 (64.22% [1858/2893])

 |  AdultDrug66 <= 22 (64.01% [1853/2895])

 JUVDRGAR <= 0 (62.07% [4436/7147])

 |  PROGTIME1 <= 17 (66.64% [1896/2845])
```

```
| ganypet33 <= 0.25 (65.42% [2883/4407])
| | PROGTIME1 <= 27.579 (68.82% [1865/2710])
| | PRIOROUTOF <= 0 (67.09% [2728/4066])
| | | PROGTIME1 <= 31 (70.37% [1926/2737])
| | | | PRIOROUTOF33 <= 0.4 (71.32% [1870/2622])
| | | age <= 16 (67.96% [1909/2809])
| | | | PRIOROUTOF33 <= 0.4 (68.89% [1847/2681])
| | | | PROGTIME1 <= 80 (68.75% [1833/2666])
```

Table 1.6

Node 2

========

Total population: 11036 instances

Target attribute: ganypet

Target value: 0 [value count in total population: 6432 instances (58.28%)]

Minimum value count for segments: 1832 instances (16.6% of total population)

Maximum branching factor: 2

Minimum improvement in target: 0.5%

ganypet=0 (58.28% [6432/11036])

PROGTIME1 <= 12 (62.92% [2026/3220])

| PROBATION <= 0 (63.62% [1833/2881])

ganypet33 <= 0.3077 (61.57% [4046/6571])

| PROGTIME1 <= 18 (65.54% [1873/2858])

| age300 > 15.936 (63.75% [1836/2880])

| | PROGTIME1 <= 164 (64.1% [1832/2858])

The next set of rules model recidivism. Again the first set requires that a rule consider at least 16.6% support and 1% accuracy. The second set requires ½% accuracy increase.

Table 1.7

Node 3

========

Total population: 11036 instances

Target attribute: ganypet

Target value: 1 [value count in total population: 4604 instances (41.72%)]

Minimum value count for segments: 1832 instances (16.6% of total population)

Maximum branching factor: 2

Minimum improvement in target: 1%

ganypet=1 (41.72% [4604/11036])

JUVDRGAR > 0 (48.68% [1893/3889])

| PRIORPERSO300 > 0.0513 (49.2% [1844/3748])

| PROGTIME1 > 3 (49.2% [1839/3738])

ganypet33 > 0.3333 (46.98% [1875/3991])

| age300 <= 16.533 (47.46% [1838/3873])

Table 1.8

```
Node 4

========

Total population: 11036 instances

Target attribute: ganypet

Target value: 1 [value count in total population: 4604 instances (41.72%)]

Minimum value count for segments: 1832 instances (16.6% of total population)

Maximum branching factor: 2

Minimum improvement in target: 0.5%


ganypet=1 (41.72% [4604/11036])

 ganypet33 > 0.3333 (46.98% [1875/3991])

 |  age300 <= 16.533 (47.46% [1838/3873])

 |  age200 <= 16.933 (47.4% [1858/3920])

 PROGTIME1 > 26 (46.97% [1897/4039])

 |  sexoff66 <= 0.3333 (47.59% [1834/3854])

 |  age300 <= 16.571 (47.5% [1834/3861])
```

The sensitivity analysis showed that 22 environmental variables were in the top 35. The sensitivity analysis determined that the most significant variables are Ganypet33, Age300, Sexoff, Progtime1, Age200, Probation and Sexoff66 (in order of significance). The regression on these variables finds that the likelihood of recidivism increases as Ganypet33, Progtime1, Age200 and Probation increase, and decreases as Age300, Sexoff, and Sexoff66 increase. Ganypet33 (recidivism ratio of other offenders within 33 meters) and Probation (kid was on probation at the time of the instant offense) are the variables that most increase the likelihood of recidivism. There are 4,815 youths that live within 33 meter rings with recidivism ratios greater than zero and 2,209 recidivated. There are 1,155 youths that were on probation at their instant offense and 569 recidivated. The recidivism rate for $Ganypet33 > 0$ is 0.4588 and $Probation = 1$ is 0.4926. The conjunction of these variables in conjunction with recidivism is:

$[(Ganypet33 > 0) \wedge (Probation = 1)] \wedge (Ganypet = 1)$

There are 555 instances where $(Ganypet33 > 0) \wedge (Probation = 1)$ and 287 recidivated. The recidivism ratio for this treatment is 0.5171, a weak but distinguishable increase over individual variables. There are 6,221 youths that live within 33 meter rings with recidivism ratios equal to zero and 3,826 did not recidivate. There are 9,881 youths that were not on probation at their instant offense and 5,846 did not recidivate. The recidivism rate for $Ganypet33 = 0$ is 0.6150 and $Probation = 0$ is 0.5916. The disjunction of these variables in conjunction with non-recidivism is:

$[(Ganypet33 = 0) \wedge (Probation = 0)] \wedge (Ganypet = 0)$

There are 5,621 instances where $(Ganypet33 = 0) \wedge (Probation = 0)$ and 3,508 did not recidivate. The non-recidivism ratio for this treatment is 0.6241.

Sexoff (instant was sex offence) was the most likely variable to decrease the likelihood of recidivism. Of the 631 sex offenders, only 162 had recidivated. This recidivism ratio of 0.2567 is much lower than the global ratio 0.4172. The C4.5 tree primarily focused on instant sex offenders because offenders in this category have a much lower recidivism rate than other types of offenders because information entropy does not consider support when choosing variables to split.

We have found at least one environmental variable, Ganypet33, that has not only contributed but in some analyses outperformed all other variables. The association analysis uses support and increase in the percentage of a target class explained by additional rules. The Ganypet33 variable is found to be significant in all association analysis experiments. As recidivism ratios of other children spatially close to a child increase, so does the child's propensity to recidivate. Other environmental variables that affect the likelihood of recidivism have been aggregated from "prior out of home placements," "prior person offenses," age of youths, and weapon and sex offenses in a child's environment.

A Preliminary Exploration of Recidivism Using Hot Spot Analysis

In order to examine juvenile recidivism in Philadelphia, we first employed hot spot analysis (here we use the term, hot spot analysis, as it is commonly used in spatial analysis). This step in the analysis was intended to help us explore the spatial nature of recidivism and develop hypotheses to be employed in subsequent analyses. In particular, we needed to know whether recidivism was concentrated in particular places or widely dispersed. Unlike a study of juvenile delinquency, we began with an understanding that those environmental factors differentiating delinquency from nondelinquency may not play as great a role in differentiating recidivists from non-recidivists. In addition to briefly presenting the methods used for conducting hot spot analysis, we also present here the findings of this analysis.

Many in the social sciences have recently recognized the necessity of spatial analysis techniques for analyzing social science data that have a locational component (Goodchild et al., 2000; Sampson et al., 2002). Certainly, problematic issues of applying non-spatial data analysis to spatial data have been well-documented (Gould, 1961), and a variety of approaches for multivariate analysis of spatial data have been developed (Anselin, 1988; Cressie, 1993; Florax and Van Der Vlist, 2003). Generally, however, these approaches have been oriented towards adapting conventional multivariate statistics to spatial data, for instance in the use of spatially autogregressive, multilevel, and random coefficient models to address spatial autocorrelation and spatial nonstationarity (Jones, 1991; Fotheringham et al., 2000; Elhorst, 2003). More recently, exploratory approaches oriented towards investigating the impact of space and place on spatial processes have gained in prominence (Anselin 1999; Fotheringham et al., 1999). For example, local measures of conventional spatial autocorrelation (e.g. Moran *I* [Moran, 1948]), summary statistics, and multivariate analysis have been developed (Anselin 1995, Ord and Getis, 1995; Boots, 2003; Brunsdon et al., 2002; Fotheringham et al., 2002).

Philadelphia is composed of five widely recognized major regions: Manayunk/Roxborough, West Philadelphia, South Philadelphia, Center City, Northern Philadelphia, and Northeast Philadelphia (Figure 1.11). These regions are divided primarily by waterways, including the Schuyllkill River, Wissahickon Creek, and Pennypack Creek. The other major regional divider within the city is the downtown area, which comprises a sixth region called Center City, and which separates South Philadelphia from Northern Philadelphia. Our initial attempt to partition program cases into sub-regions consisted of simply assigning each case to one of the six major regions within which the case resided.

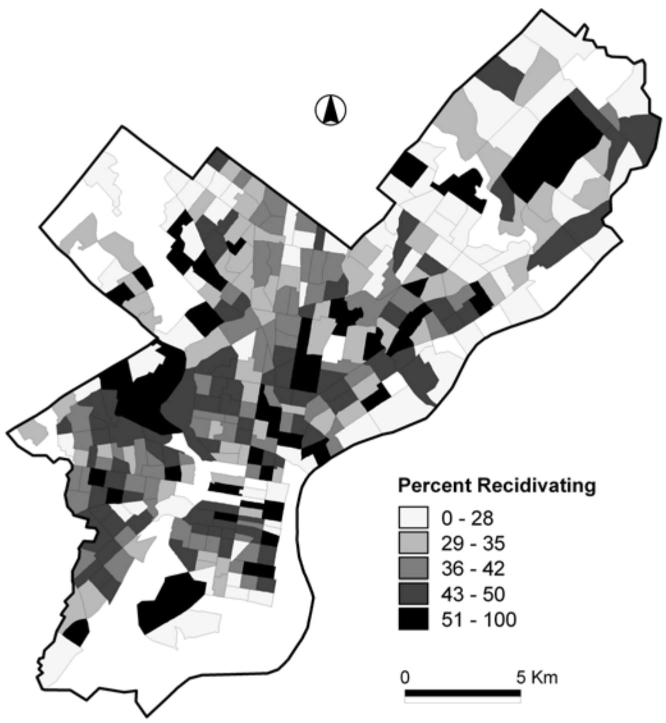
The recidivism ratio was then calculated for each case. The recidivism ratio is the proportion of all cases that recidivated, defined for each case as whether the juvenile in that case re-offended within six months of completing the court-ordered program. The recidivism ratio is calculated for a set of cases over an area, and was calculated in the following way. First, a statistical surface of the density of cases was generated using a 1 km bandwidth over a 100m resolution grid, so that each 100 m by 100 m square grid cell encoded the number of cases per square kilometer as counted over an area extending 1 km from the center of each cell. This point density operation is a well-established technique for point pattern analysis (Bailey and Gatrell, 1995).

Figure 1.11


Somerton
Northeast
Torresdale
Roxbo-
rough
German-
town
Oak Lane
Oxford
Circle
Man-
ayunk
Logan
Frankford
Northern
Hunting
Park
Kensington
Wynne-
field
Strawberry
Mansion
Over-
brook
West
Center City
Cobbs
Creek
Point Breeze
0
5 Km
King-
sessing
Girard
Estates
Penns-
port
Eastwick
South

The density surface calculation was done individually for each region, so that the density calculation did not incorporate cases located across region boundaries. Our motivation for isolating regions in this regard stems from the belief that juvenile recidivism behavior is unlikely to be influenced by characteristics of an area that may be within 1 km of the juvenile, but are separated from the juvenile by a major river or other large physical barrier. Because the case density is far sparser in Northeast Philadelphia as compared to the other regions, a 1.5 km bandwidth was used for this region only. The resulting statistical surface, referred to as the 'case density surface,' is shown in Figure 1.12.

Figure 1.12



We had a number of reasons for parameterizing point density function using a 100 m resolution grid and 1 km and 1.5 km bandwidths. A spatial resolution of 100 m was chosen as a way to both maintain a fine enough grid to capture the spatial heterogeneity in recidivism ratio while also maintaining a data volume small enough for efficient computation on a desktop computer. We acknowledge that the choice of a 1 km bandwidth is perhaps somewhat arbitrary. We experimented with a number of bandwidth options, including the use of a number of nearest neighbors and a bandwidth that expanded up to 3km in areas of sparse cases and contracted to 1 km in areas of denser cases. However, we ultimately decided to use a relatively small fixed bandwidth of 1 km (and 1.5 km in Northeast Philadelphia), as this radius captures the immediate area around each location, while still capturing a sufficient number of cases for calculating the recidivism ratio for most grid cells.

An operation analogous to the generation of the case density surface was performed only for those cases which had recidivated, yielding a density surface of recidivism for each region (Figure 1.13). We refer to this surface as the 'recidivism density surface.' Again, a 1 km bandwidth was used to calculate the recidivism density surface, except for Northeast Philadelphia, where a 1.5 km bandwidth was used. A 'recidivism ratio surface' was derived by dividing the case density surface by the recidivism density surface, so that each cell encoded the proportion of cases which had recidivated in the proximity of that cell (Figure 1.14). The case density and recidivism ratio values were encoded for each case by retrieving the values for the respective statistical surfaces for each case location.

A local cluster analysis was employed to classify cases into sub-regions of similar recidivism ratio. This part of the analysis was performed on a set of points derived from the recidivism ratio surface, where those cells from the recidivism ratio surface that contained one or more cases were converted to points (derived from the geometric center of the grid cell). We refer to these points as 'recidivism ratio points.' Note that some grid cells contained multiple cases, but most grid cells contained zero cases. Our motivation for performing the local cluster analysis on these recidivism ratio points, as opposed to, say, all the case locations, is based on a conceptualization of the recidivism ratio variable as a surface that varies continuously over space. We consider the grid cells at the case locations to be samples of this continuous surface, which we intend to analyze for characteristics of local spatial dependency.

Figure 1.13

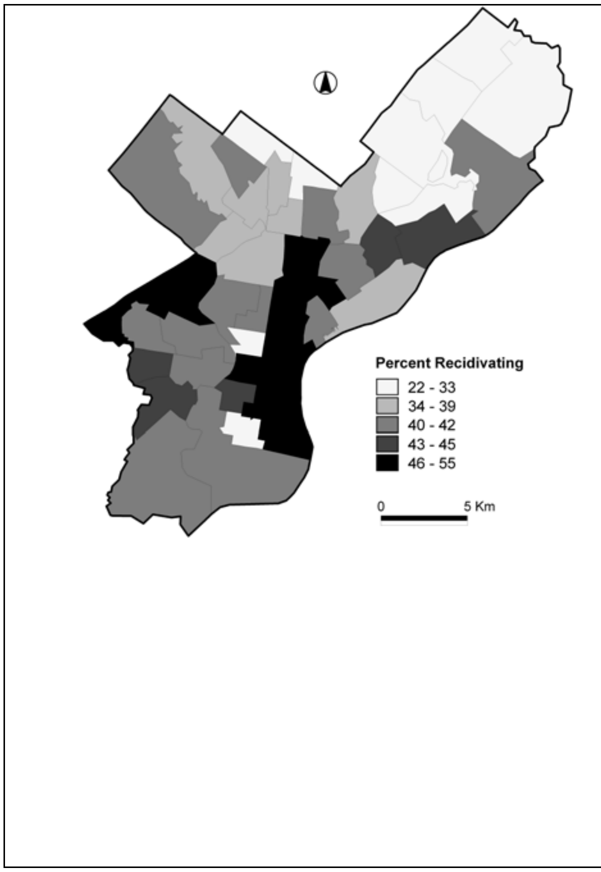
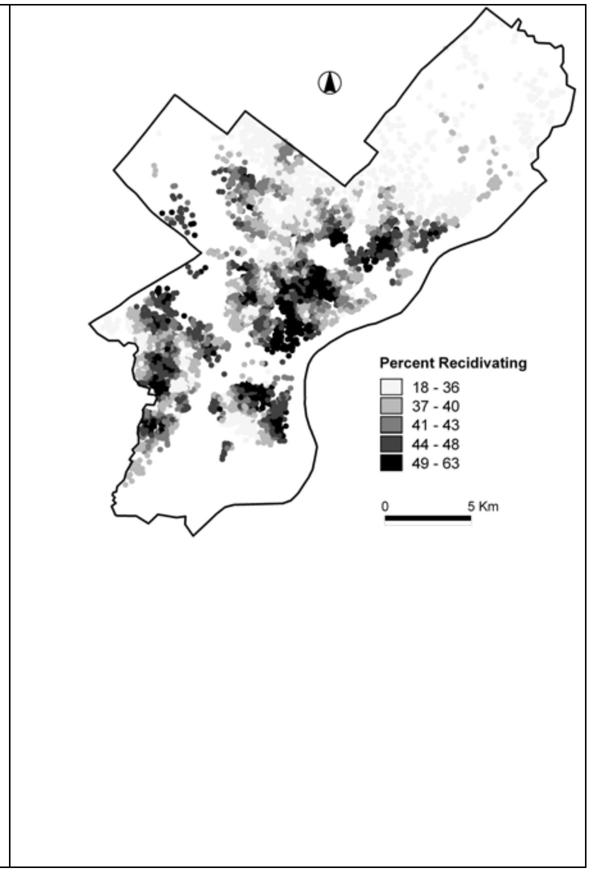


Figure 1.14



We also masked those portions of the recidivism ratio surface in which the ratio was calculated using less than ten cases (i.e. less than 10 cases in the denominator of the ratio calculation), as there is little meaning in a ratio calculated over such a sparsely populated area – a change in recidivism in just one or two cases can have a dramatic effect on the recidivism ratio calculation when there are less than ten cases total. We note that the vast majority (93%) of recidivism ratio points used in the local cluster analysis were calculated using more than 50 cases in the denominator. This extraction resulted in a point data set with 5,608 observations (Figure 6). This reduction in the number of observations from the raw number of cases (n=11,659) also aided in the computational performance of the local cluster analysis, as such an analysis is computationally intensive.

Prior to the local cluster analysis, the well-known Moran's *I* statistic was applied to the recidivism ratio points to confirm that the recidivism ratio is indeed spatially clustered (I=0.77, significance < 0.01). The Getis G_i* statistic (Ord and Getis, 1992; Getis and Ord, 1995) was then used for the local cluster analysis. This statistic captures significant differences between a local neighborhood and the global mean for a particular variable, and is expressed as

$$G_i^*(d) = \frac{\sum_j w_{ij}(d) z_j - W_i^* \bar{z}}{s\{[(nS_{1i}^*) - W_i^{*2}]/(n-1)\}^{1/2}}$$

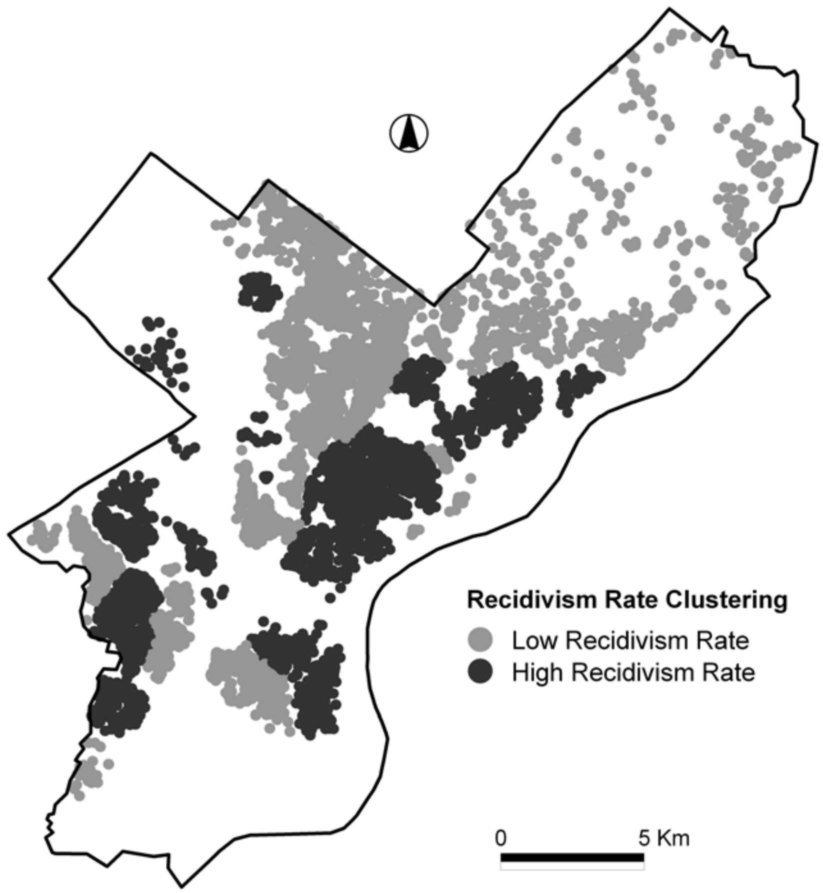
where $w_{ij(d)} = 1$ if location *i* is within distance *d* of location *j*, and $w_{ij(d)} = 0$ if it is not, and $\bar{z}$ and s^2 denote the sample mean and variance.

In the present case, Getis G_i^* was used to identify, for each recidivism ratio point, whether the recidivism ratio in the immediate neighborhood around that point was significantly different from the recidivism ratio of the entire data set. Thus, unlike with the generation of the case and recidivism density surfaces, which was performed individually for each region, the local cluster analysis was performed once over the entire Philadelphia data set. Parameterization of the G_i^* calculation demands specification of the spatial weights matrix. We experimented with a number of specifications including the use of inverse distance and inverse distance squared weighting schemes. Because the density of points varied greatly throughout the data set, these distance weighted spatial weights matrix specifications tended to skew the results for somewhat isolated points or points on the edge of the region, because in these situations the calculation of the G_i^* statistic may be based almost exclusively on one or two points located nearby the focal point. Ultimately, we decided to use a 1 km fixed bandwidth, so that every point pair within a 1 km distance of each other was denoted with a '1,' and given equal weight in the G_i^* calculation, and all other point pairs were denoted with a '0.'

Following the calculation of the G_i^* statistic for each recidivism ratio point, each original case point was assigned a G_i^* value based on the case point's spatial association with a recidivism ratio surface point. Each case was then mapped according to its G_i^* value using a four-class scheme: 1) significantly high recidivism ratio cluster, 2) significantly low recidivism ratio cluster, 3) not significant cluster, and 4) G_i^* not calculated because there were less than 10 cases with which to calculate the recidivism ratio. A significance threshold of 90% was employed, which is slightly more relaxed than the 95% threshold typically used in the social sciences. However, as Figure 7 shows, the clusters of high and low recidivism ratio tend to be separated by 'transition zones' where the G_i^* value is not significant. A higher significance threshold merely serves to expand these transition zones by incorporating more cases into the 'not significant' class. As our aim is to identify broadly-defined areas of high and low recidivism via an exhaustive (to the degree possible) partitioning of cases, we employ a relatively relaxed significance threshold of 90%.

This map was displayed overlain with streets, railroads, and waterways in order to visually and manually identify spatially coherent clusters of cases (Figure 1.15). Cases that belonged to the same mapped class (e.g. high and low recidivism ratio clusters) and that are not separated by major physical barriers, including the boundaries separating the six major regions of the city, were considered members of the same sub-region. All cases were exhaustively classified into sub-regions in this manner.

Figure 1.15

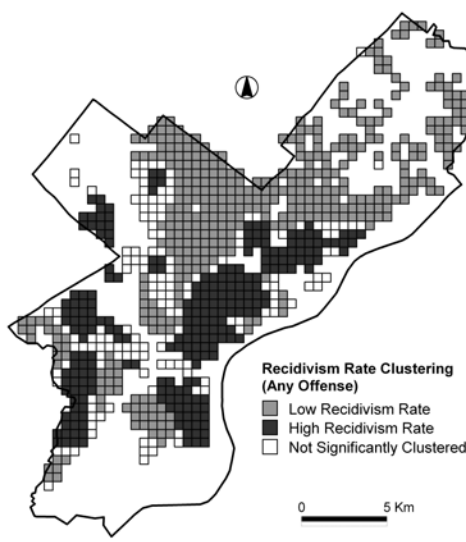


Results

This manual grouping of cases based on both the the G_i* statistic and physical barriers yielded a partitioning of all cases into 66 individual sub-regions (Figure 1.16). The largest sub-region, in central Northern Philadelphia, contained 2,835 cases. The sub-region median for the number of cases was 20, as many of the sub-regions represented isolated areas or residential 'peninsulas' with just a few cases that typically were either not significantly clustered or did not receive a recidivism ratio variable because there were less than 10 cases within the immediate proximity. The juxtaposition of low and high recidivism areas spatial nonstationarity of recidivism, a topic we take up in more detail in the next five chapters.

This analysis also provided the first indication that type of offense is spatially dependent. Table 1.17 shows a pattern for person offenses that is markedly different from that of recidivism as a whole. In the remaining chapters, we continue this line of inquiry, conducting several analyses that focus on type of recidivism offense.

Figure 1.16



CHAPTER 2
NEURAL NETWORK ANALYSIS OF JUVENILE RECIDIVISM

INTRODUCTION

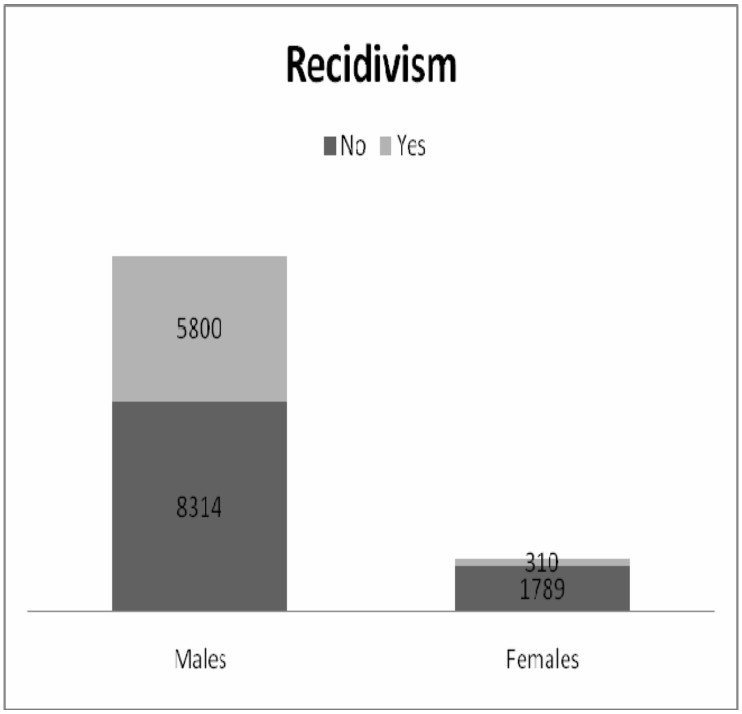
Our research is focused on examining juvenile justice data using data mining techniques to predict and model patterns of juvenile recidivism. Juvenile recidivism is an insufficiently explored domain within the data mining community. This is in part due to restricted access to juvenile records and also because of the difficulties that are inherent to applying data mining techniques to social sciences data. This data is difficult to model because variables aimed at describing human behavior are often high noise, low signal and almost random in large part due to unpredictable aspects of human behavior. Large social science datasets typically have large amounts of missing values. These studies are often performed over long periods of time and over large regions of space. These data can be aggregated from many different sources of origin.

In our study observations are a mix of case data about the child and the child's family, the treatment program that the child attended, and spatially generated environmental data about the child's neighborhood. We have investigated various data mining techniques to explore to what extent we can increase the accuracy of both modeling (accuracy on training data for the phenomenon understanding) and prediction of recidivism (accuracy on unseen cases for possible preventive purposes). Our hypothesis is that accuracy of modeling and prediction can be improved by partitioning the data into smaller, more homogeneous sets followed by development of specialized models for identified groups.

THE DECISION TO REMOVE FEMALES FROM THE STUDY

We initially performed a cursory analysis on the ProDES data set to try to understand some of its basic characteristics. We tested a subset of the variables with various attribute selection techniques, ran various classification models and scrutinized the results. One important finding was that male and female offenders should not be analyzed in the same classification experiments. In the first subset of ProDES that we analyzed, there were 14,114 male and 2099 female cases. Of the males, 5,800 had new petitions prior to 6 months from program release, which is a recidivism rate of 41.09%. Only 310 of females recidivated: a 14.77% recidivism rate. These subgroups have very different characteristics with respect to recidivism. At first glance the "sex" variable seems to be a significant predictor for our problem. While it's truly significant for predicting recidivism, it has been found to be detrimental to our models. We observed the actual predictions and found that the classification models were finding that no females recidivated. Our concern is that while models with females had higher overall accuracy, the power of this variable may mask valuable information contained in much weaker predictors. This finding led us to the conclusion that females should be removed and analyzed separately. The figure below shows the differences between male and female cases.

Figure 2.1.

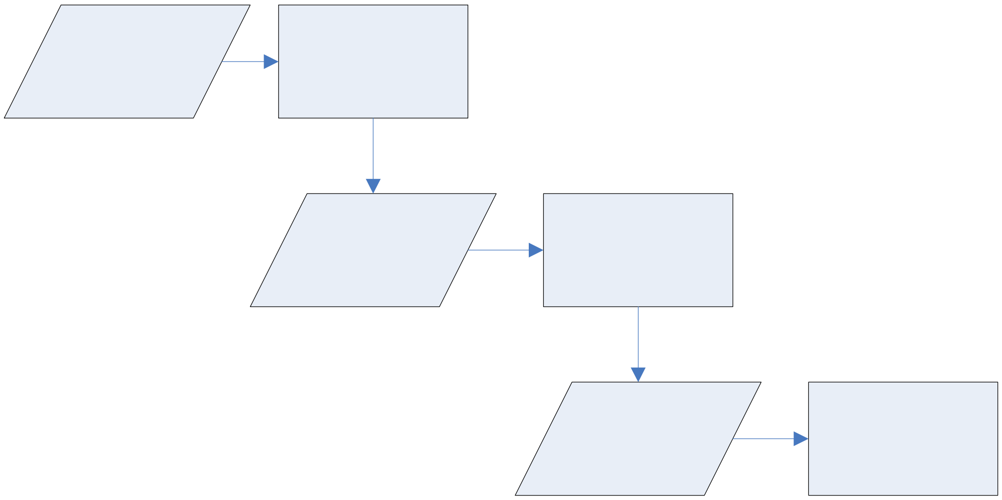


Methodology

The flow chart below shows an overview of the method. The first process is the partitioning of the original data set using Plaid [Lazzeroni and Owen]. Next, the Chi-Squared Measure [Shaomin and Flach] is applied. The partitions and the original data set (used as a baseline) are predicted and modeled with neural networks, after which the results of the partitions are compared to those of the original set.

The Plaid method is discussed in Chapter 1, under Analytic Methods, and is utilized again in Chapter 3, where the spatial patterns of layers produced by co-clustering are analyzed. In this chapter, we explored the partitioned data to determine if prediction of recidivism could be improved over that found for the data set as a whole.

Figure 2.2.



Data Partitioning

The Plaid algorithm was used as a method of data partitioning. Our method uses Plaid to decrease noise and increase signal by producing bi-cluster (clustering both rows and columns of the data matrix) type layers consisting of statistically similar cases in a given subset of attributes and excluding cases that are statistically distracting to the predictions and models. Plaid samples the input data with replacement, which means that both cases and variables can be present in multiple layers. This approach differs from traditional clustering methods that sample without replacement. Traditional clustering partitions data under the concept that all data points in a cluster share some similarity with all other points within the same cluster and are dissimilar with all points outside of the cluster.

Attribute Selection

The Chi-Squared Measure is a heuristic attribute selection technique based on a contingency table [Pearson] and the Chi Squared test [Lehmann and Romano]. This method evaluates the measure the value of an attribute with respect to target class. It is assumed that the attribute and the target under consideration are independent and test this assumption with the Chi Squared measure. There are two target classes, P (positive) and N (negative), and r attribute classes, $C_1 \ldots C_r$. The contingency table is defined as:

Table 2.1.

	Target = P	Target = N	Total
Attribute = C_1	$n_{1P}\left(\mu_{1P}\right)$	$n_{1N}\left(\mu_{1N}\right)$	n_{1*}
...	...	...	...
Attribute = C_r	$n_{rP}\left(\mu_{rP}\right)$	$n_{rN}\left(\mu_{rN}\right)$	n_{r*}
Total	n_{*P}	n_{*N}	n

n_{ij} is the number of cases for which the value of the attribute is C_i and the value of the target is j,

$n_{*j} = \sum_{i=1}^{r} n_{ij}$, $n_{i*} = n_{iP} + n_{iN}$, $n = n_{*P} + n_{*N}$, and $\mu_{ij} = \frac{n_{*j} n_{i*}}{n}$, where $i = 1, \ldots, r$ and $j = P, N$.

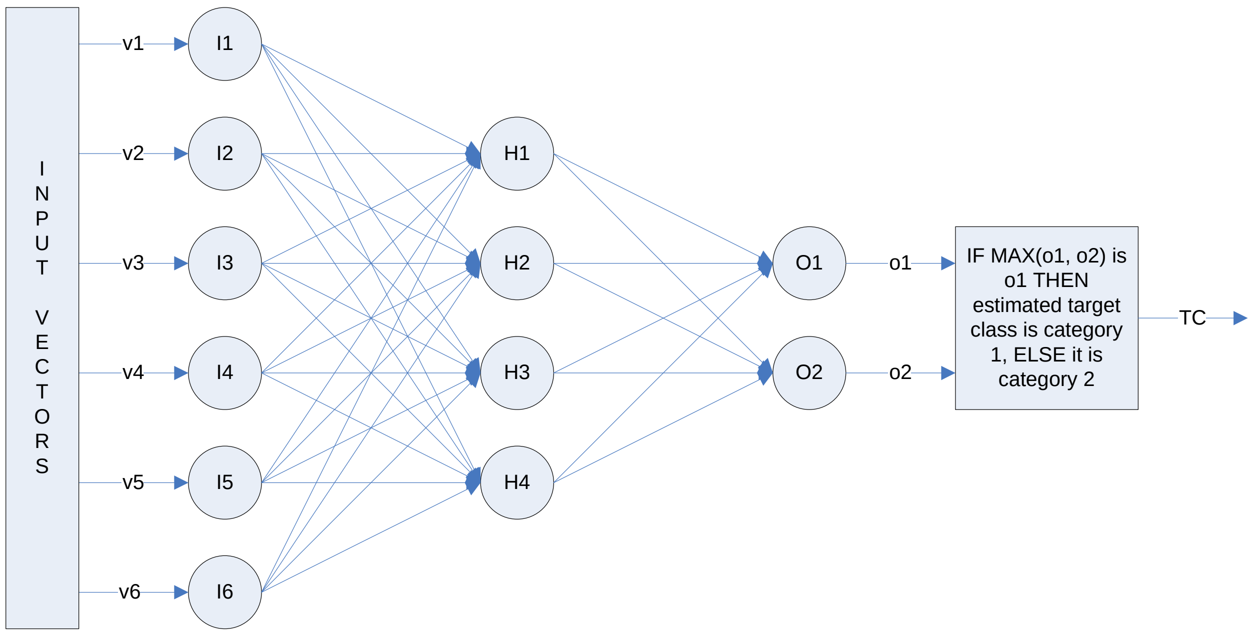
The Chi-Squared measure is defined as:

$$\chi^2 = \sum_{i=1}^{r} \left(\frac{\left(n_{iP} - u_{iP}\right)^2}{u_{iP}} + \frac{\left(n_{iN} - u_{iN}\right)^2}{u_{iN}} \right) \qquad (1)$$

Neural Networks

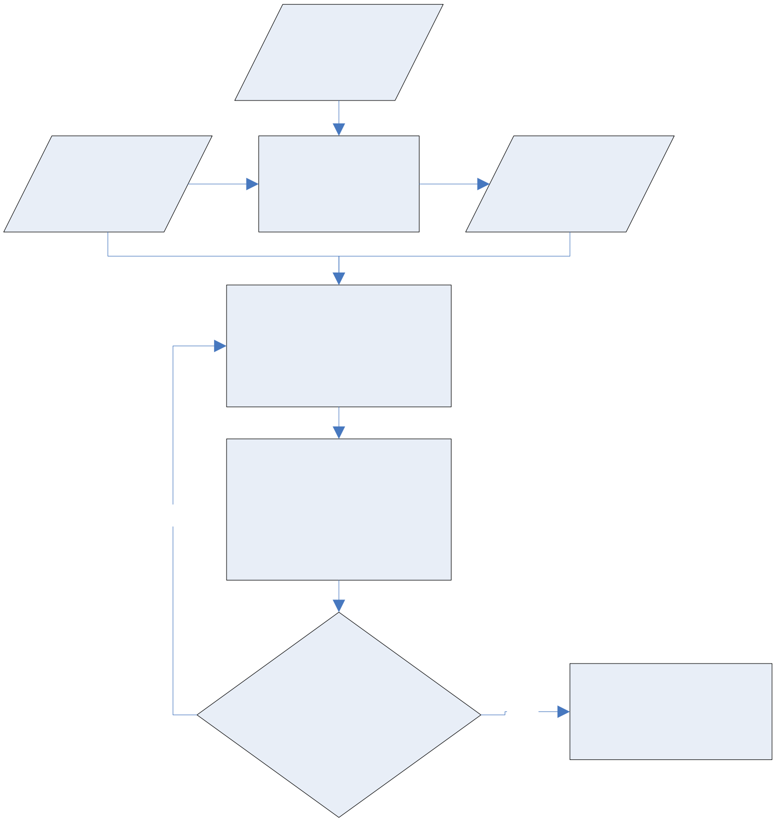
Our modeling and prediction experiments were based on artificial neural networks. These nonlinear models are universal function approximators, , as proven by the Cybenko theorem [Cybenko], which are often able to model dependencies among attributes and a response that are too complex for logistic regression or decision trees. The feedforward type used in our experiments has an input layer of nodes where observed attributes are applied, an output layer, where the predictions are rendered and another hidden layer of nodes with nonlinear activations fully connected between the input and the output nodes with weighted directed links. An example feedforward network with four attributes in the input layer, a hidden layer with six hidden nodes and an output layer with three output nodes is shown at Figure [x]. The data fed into this network would have four variables and three target classes.

Figure 2.3.



For given weight assignments on the interconnection pattern within the network, each processing node computes a nonlinear function of its weighted input sum, which was the hyperbolic tangent in our experiments that scales the sum from (-∞, ∞) to (-1, 1). Weights in a neural network are initialized randomly and are adapted through a gradient-descent optimization as to minimize the mean square error on training examples. The flowchart below shows an abbreviated process for the training and testing operations for typical neural networks.

Figure 2.4.



The back propagation algorithm used for optimizing weight parameters is based on an efficient computation of partial derivatives of an approximation function realized by the network []. In this training process, a set of training instances is repeatedly fed through the network through possibly hundreds or thousands of iterations, calculating an output and adjusting the weights based on their estimated influence on the observed errors on a training example. The error correction function takes the partial derivative of the weight matrix, to find minima for the instances' outputs when compared to the target values, and adjusts the weight and bias to calculate this derived value. This gradient descent optimization is aimed at reducing the distance between the network's estimate and the actual target value. The effect is gradual and should improve with subsequent iterations until it converges on an optimal set of estimates for the inputs' target classes. Error functions have momentum and learning rate values to control the adverse effects of erratic instances and local minima. The learning rate controls the amount of change that can occur in any given correction. The momentum preserves the overall direction of adjustment, giving consideration to past corrections.

Experiments

The Data/Preprocessing

The current complete data set has 6675 cases, all male cases that were assigned to residential programs, and 436 variables of categorical, continuous, ordinal and binary type. The target variables for the analysis are drug recidivism (XDrugs), person recidivism (XPerson), property recidivism (XProperty) and all recidivism (ganypet). These are all binary targets. This data had [x] missing fields, which is approximately [y]%. Variables with more than 30 percent missing data and cases with more than 20 percent were removed from the set before data replacement. The reduced data set has 6675 cases. 174,056 of 2,963,700 fields or slightly more than 5.87% of the data set were marked as missing data. We have replaced missing data with our novel cluster-based technique that replaces missing fields based on similar cases (for details see attached manuscript "Dynamic Clustering-Based Estimation of Missing Values in Mixed Type Data." by Ayuyev, V., Jupin, J., Harris, P. and Obradovic, Z.). This method has been proven to outperform Mean-Mode based missing data replacement. The 436 initial variables were expanded to 839 after all of the categorical attributes were converted to the corresponding bit vector representation (e.g. to eliminate bias an attribute with three category is represented as three binary attributes where 100, 010 and 001 values correspond to the first, second and third category).

Data Partitioning

Plaid algorithm was used as a method of data partitioning for the follow up neural network modeling. Our hypothesis was that using Plaid will decrease noise and increase signal by producing bi-cluster (clustering both rows and columns of the data matrix) type layers consisting of statistically similar cases in a given subset of attributes and exclude cases that are statistically distracting to the predictions and models. Plaid samples the input data with replacement, which means that both cases and variables can be in multiple layers.

The variables selected for inputs to Plaid are listed in Table 2.2 below. These variables were selected because they have demonstrated some value for recidivism analysis in previous experiments. The data that was analyzed with Plaid did not contain any of the recidivism target variables.

Table 2.2.

Variable	Definition
sexoff	Any charges for sexual offense
White	Juveniles race is white
Hispanic	Juveniles race is Hispanic

Probation	Supervision level at instance: probation
ParentalCrim	Parent has criminal history
LiveInstitution	Living arrangement: institution
PriorPerson	Youth had prior person offence
Juvdrgar	Youth had prior drug arrests
Prioroutofhomepl	Youth had prior out of home placements
InstantPerson	Instant offense was person related
InstantProperty	Instant offense was property related
sibarr	Any siblings arrested
jhismh	Any history of mental health problems
victinj	Injury to victim
age	Age at instant
AlcoholAbuse	Any history of alcohol abuse for youth
DrugAbuse	Any history of drug abuse for youth
ParenDeceased	At least one parent is deceased
den_dr_sale	Density of drug sales near youths home
den_person	Population density near youths home
p_vacant	Percent vacant housing in census block
p_black	Percent African American in census block
p_spanish	Percent of census block that are Hispanic
p_highsch	Percent of census block that graduated high school
kcnt_1km	Number of youths within 1 kilometer
gi	Recidivism density excluding this juvenile
ParSubAbuse	Parent has drug or alcohol problem

The Plaid analysis produced 52 layers. We selected layers from the Plaid results that had at least 400 cases and no more than 4 variables. Layers with less than 400 cases are not likely to be very useful and layers with more than 4 variables would be difficult to analyze. There were 13 layers that met these conditions. These layers have much fewer cases than the input data. It has been observed, during preliminary trials, that decreasing the number of cases is detrimental to both prediction and modeling as such layers can potentially reveal some interesting information about partitions of the data. The youths in these layers have some characteristics in common with each other but the recidivism rates of some partitions differ greatly from the whole data set. To evaluate the discovered partitions versus random selection, we compared the Plaid layer based partitions versus 1000 random partitions of the complete data set to 52 subsets that had the same number of cases as the corresponding Plaid layers. In addition, to explore the possibility that Plaid layers are determined by their dominant attribute(s) we compared the obtained partitioning versus equal size random sets selected according to the dominant attribute associated the corresponding Plaid layer.

Attribute Selection

We created data subsets for each of the 13 layers that contain all attributes columns in the complete data set but only the cases that were included in each of the layers. Our intention was to identify the significant predictors for each layer from all possible candidates. The Chi-Squared measure [Shaomin and Flach] was calculated for all of the attributes within each layer with respect to all four target variables: drug, person, property and all recidivism types. The Chi-squared measure is a heuristic attribute selection technique based on a contingency table [Pearson] and the Chi Squared test [Lehmann]. This method evaluates the value of an attribute with respect to target class on an individual basis.

The variables were sorted in descending order by the measure and the best candidates were kept for each layer-recidivism type set. Up to 40 variables were picked for each set. However, some layers had less than 40 because all variables with a Chi-Squared measure of zero were excluded. This procedure is intended to remove the least useful variables from the set. We also used the Chi-Squared measure to select the top 40 attributes from all cases in the complete data set to use as the basis for our comparison. When completed, we had 56 data sets, 14 for each recidivism type.

Neural Network Analysis

We created one neural network prediction and one model for each of the 56 data sets, or 112 trials. The networks were configured with the following parameters:

Layers: 3

Hidden nodes: Calculated as (input nodes – output nodes)/2

Learning rate: 0.3

Momentum: 0.2

Decay: Yes

Normalized attributes: Yes

Training time: 2000

This configuration is similar to Figure 2.3 above. It has 3 layers: an input layer, a hidden layer and an output layer. The structural differences are in the number of input, hidden and output nodes within the networks layers. There are up to 40 input nodes for each neural network, one for each input variable. The number of hidden nodes is calculated based on the number of inputs and outputs. A neural network with 40 inputs and 2 outputs will have 21 hidden nodes. Too many hidden nodes can lead to over fit classifiers. There are two output or target nodes. One node represents positive recidivism and the other represents negative recidivism. The winning target node will be the node with the highest value after the input is calculated within the network. The learning rate determines how much the connection weights are allowed to change during each iteration of the neural network's error correction function. The momentum decreases the potential for thrashing or drastic changes in the connection weights within the network. It takes into consideration the previous corrections by nudging the weights towards a single direction to avoid erratic updates. The decay parameter slightly decreases the learning rate of the network with each iteration. This will stop the neural network from diverging from the target and increase the network's performance. The normalized attribute parameter scales the input values between -1 and 1. This will ensure that inputs with very high magnitudes will not dominate the model. The training time is the maximum number of iterations performed by the normal network. An iteration is a complete pass over the input data set, after which the error correction function updates the connection weights.

RESULTS

Data Partitioning

Layer 1

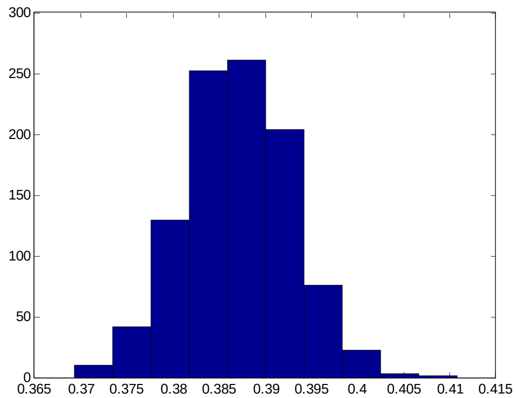
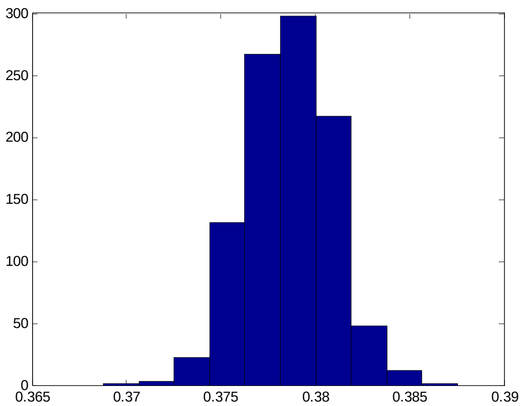
Layer 1's partitioning was based on three variables: age of juvenile at instance (age), percent of census block that graduated high school (p_highsch), and percent of census block occupants that are African American (p_black). This layer has 3357 cases and a recidivism rate of 35.24%. This layer has retained variables that pertain to primarily spatial demographic data. The table below shows the mean for these variables within the layer and for the complete set. The LMean column is the mean for each of the variables in the layer and the GMean is the global mean from the complete data set. The mean age is nearly the same, the percent of high school graduates in tract is slightly higher but percent African Americans in tract is very much higher.

Table 2.3.

	LMean	GMean
age	15.55	15.64
p_highsch	62.10	58.90
p_black	90.34	64.66

The distribution in Figure 2.5 below shows the results from the random selections of 1000 subsets of ProDES with the same number of cases as Layer 1. Notice that the recidivism rate for Layer 1 (35.24%) is significantly lower than the average of the random subsets. Since the recidivism rate for the random selection is 38.7% and the standard deviation is 0.058%, Layer 1 is 5.97 standard deviations lower than the randomized norm. The distribution on the right shows the random selection of tracts in the data set with 75% or more African American occupants. The 75% value was selected because it is one standard of deviation less than the Layer 1 average for percent African Americans in tract. In this case, the difference between Layer 1 and the average of random subsets is 11.93 standard deviations. Both of these experiments make a strong case to support the argument that this layer is not based on a random function. The distribution on the right suggests that it is the combination of the variables included in (and possibly those excluded from) this layer that have an effect on the recidivism rate, not just the dominant variable.

Figure 2.5.: Layer 1

Random selection:	p_black > 75% (mean = 90.34% std = 14.63%)
	
Mean: 0.3870	Mean: 0.3786
Std: 0.0058	Std: 0.0023

Importantly, this layer demonstrates a theme of the importance of education. As can be seen from Figure 2.5, when census blocks with populations greater than 75% African American are selected, the recidivism rate is 37.9%. This layer shows that when the African American race is combined with a higher than average proportion of the population with high school educations, recidivism is lower than average (35.2%).

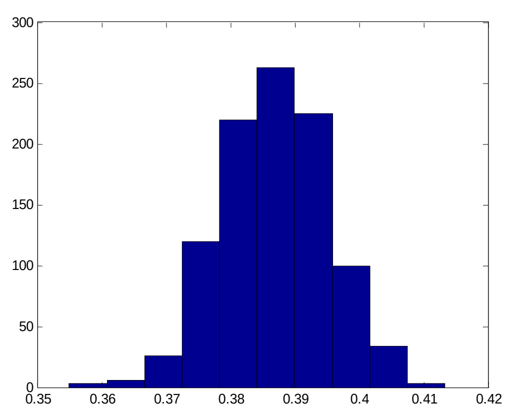
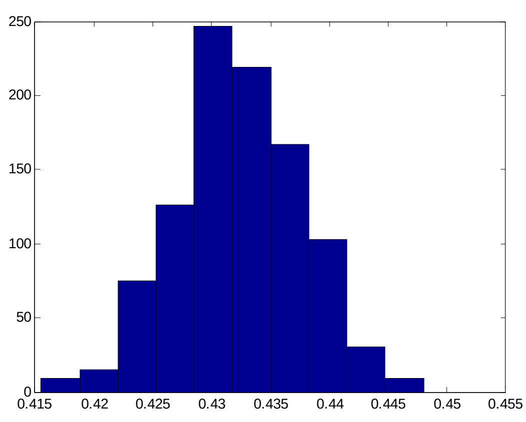
Layer 3

Layer 3 has a recidivism rate of 42.74%, 2279 cases and 2 variables: age of juvenile at instance (age) and juvenile history of drug abuse (DrugAbuse). The table has two additional columns because juvenile history of drug abuse is binary. LCntB is the count of positive binary values in the set for a binary variable and GCntB is the number of positives in the complete set of cases. This layer is completely comprised of youths that have had drug abuse problems. The distributions below show the random selections from the complete set of cases (right) and only cases where the youth has a history of drug abuse. This recidivism rate for this layer is heavily influenced by the drug abuse variable. However, not all drug abuse cases from the complete set were selected, which suggests that the age at the point of disposition variable and the excluded variables have influence on the selection of cases for this set. This layer's recidivism rate is 4.94 standard deviations higher than the norm from a randomized subset from the complete set and 0.94 lower than the rate for all cases where the youth has drug abuse problems.

Table 2.4.

	LMean	LCntB	GMean	GCntB
age	16.24		15.64	
DrugAbuse	1.00	2279	0.4634	3136

Figure 2.6.: Layer 3

Random selection:	Select cases where DrugAbuse = 1
	
Mean: 0.3869	Mean: 0.4324
Std: 0.0082	Std: 0.0053

Comparing the mean recidivism rate of youths with drug abuse histories (.43) and the rate for Layer 3 demonstrates that DrugAbuse contributes more to the Layer 3 rate than does age.

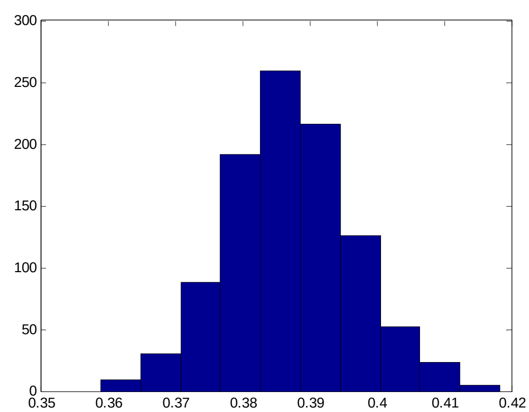
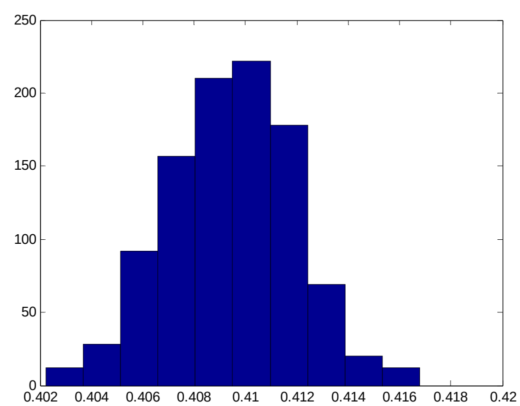
Layer 5

Layer 5 has a recidivism rate of 40.07%, 1989 cases and 2 attributes: youth lived in an institution at instant (LiveInstitution) and instant offense was property (InstantProperty). Again there is a dominant variable but some cases were excluded by Plaid. All the youths in this layer had committed property crimes as their instant offense. This layer's recidivism rate is 1.44 standard deviations higher than the norm from a randomized subset from the complete set and 3.35 lower than the rate for all cases where the youth's instant offense was property related.

Table 2.5.

	LMean	LCntB	GMean	GCntB
LiveInstitution	0.3338	664	0.3218	2178
InstantProperty	1.00	1989	0.3104	2101

Figure 2.7.: Layer 5

Random selection:	Select cases where InstantProperty = 1
	
Mean: 0.3872	Mean: 0.4094
Std: 0.0094	Std: 0.0026

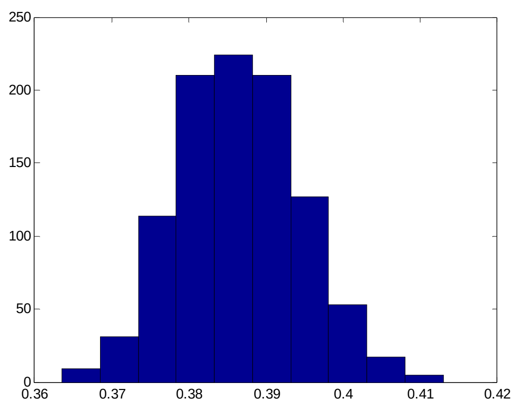
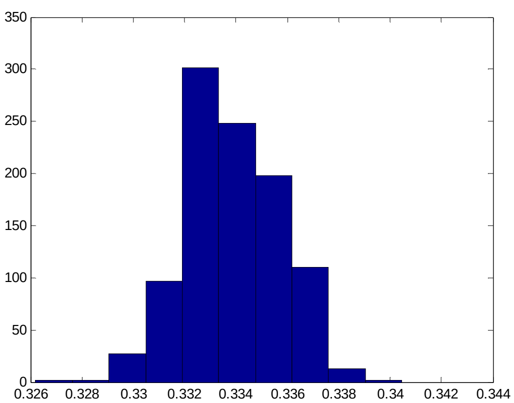
Layer 6

Layer 6 has a recidivism rate of 33.42%, 2385 cases and 2 attributes: victim injured at instant (victinj) and instant offense was person (InstantPerson). All the youths in this layer committed person offenses as their instant offense. This layer's recidivism rate is 6.54 standard deviations lower than the norm from a randomized subset from the complete set and 0.17 higher than the rate for all cases where the youth's instant offense was person related.

Table 2.6.

	LMean	LCntB	GMean	GCntB
victinj	0.4553	1086	0.1674	1133
InstantPerson	1.00	2385	0.3647	2468

Figure 2.8.: Layer 6

<table>
<tr><td>Random selection:</td><td>Select cases where InstantPerson = 1</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Mean: 0.3865</td><td>Mean: 0.3339</td></tr>
<tr><td>Std: 0.0080</td><td>Std: 0.0018</td></tr>
</table>

Not surprisingly, youths who have committed person offenses have lower than average recidivism rates.

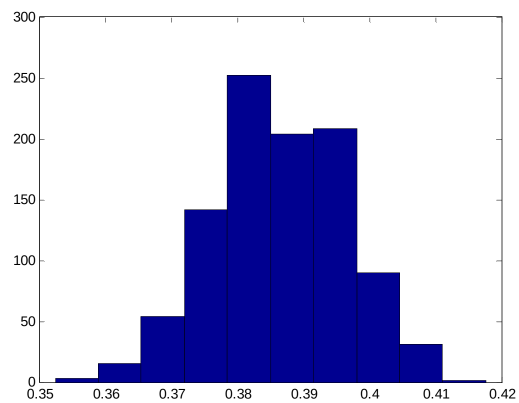
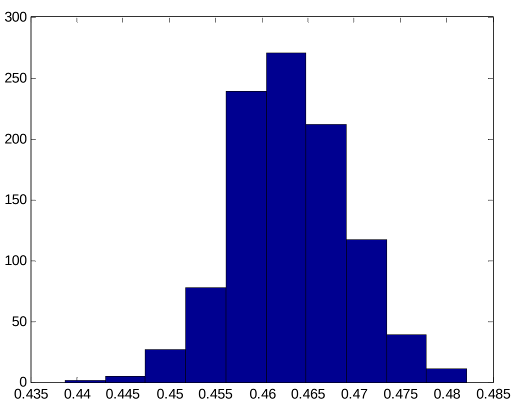
Layer 8

Layer 8 has a recidivism rate of 46.05%, 1657 cases and 3 attributes: number of youths within 1 kilometer of the youth's primary residence at instant (kcnt_1km), the recidivism density in the youth's census block (gi) and youth has previous drug arrests (Juvdrgar). All these youths had prior drug arrests. This layer's recidivism rate is 7.47 standard deviations higher than the norm from a randomized subset from the complete set and 1.02 lower than the rate for all cases where the youth has had prior drug arrests.

Table 2.7.

	LMean	LCntB	GMean	GCntB
kcnt_1km	213.16		185.75	
gi	4.1517		1.6324	
Juvdrgar	1.00	1657	0.3212	2174

Figure 2.9: Layer 8

<table>
<tr><td>Random selection:</td><td>Select cases where Juvdrgar = 1</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Mean: 0.3865</td><td>Mean: 0.4632</td></tr>
<tr><td>Std: 0.0099</td><td>Std: 0.0062</td></tr>
</table>

This layer is quite different from Layer 6. The concentration of recidivating youtns and involvement in drug offending is associated with a high rate of recidivism (.46).

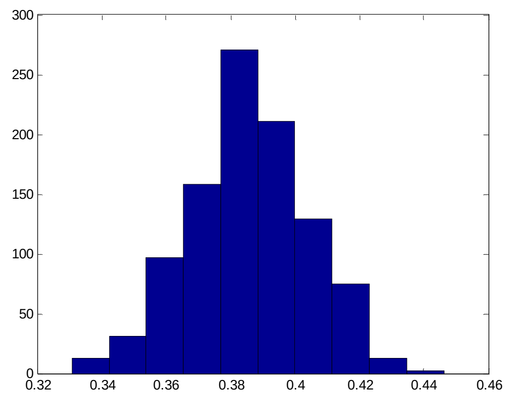
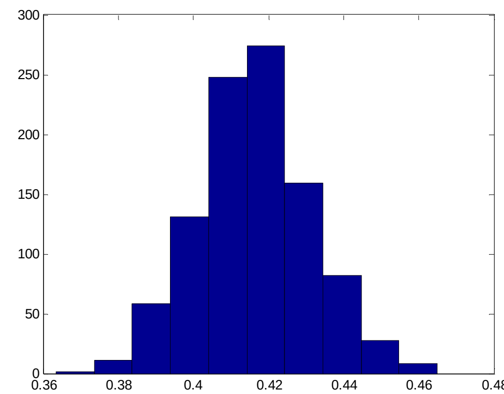
Layer 15

Layer 15 has a recidivism rate of 41.03%, 641 cases and 2 attributes: youth has had an alcohol problem (AlcoholAbuse) and parent has had a substance abuse problem (ParSubAbuse). All the youths in this layer have at least one parent that has had a substance abuse problem. This layer's recidivism rate is 1.31 standard deviations higher than the norm from a randomized subset from the complete set and 0.38 lower than the rate for all cases where the youth's parents have had substance abuse problems.

Table 2.8.

	LMean	LCntB	GMean	GCntB
AlcoholAbuse	0.2652	170	0.2114	1431
ParSubAbuse	1.00	641	0.2370	1604

Figure 2.10.: Layer 15

<table>
<tr><td>Random selection:</td><td>Select cases where ParSubAbuse = 1</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Mean: 0.3859</td><td>Mean: 0.4161</td></tr>
<tr><td>Std: 0.0185</td><td>Std: 0.0153</td></tr>
</table>

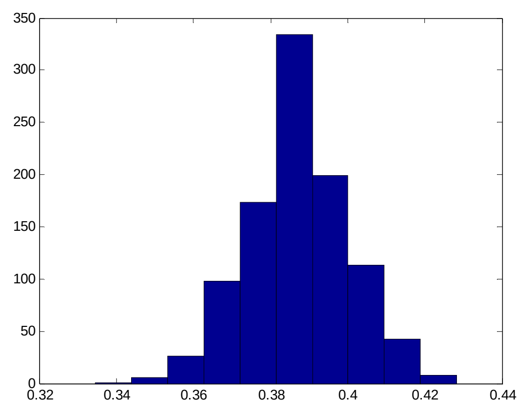
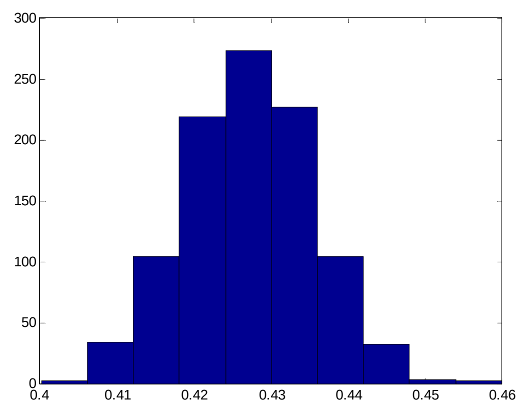
Layer 18

Layer 18 has a recidivism rate of 38.82%, 1172 cases and 1 attribute: one of the youth's siblings has an arrest history (sibarr). All the youths in this layer have at least one sibling that has an arrest history. This layer's recidivism rate is 0.08 standard deviations higher than the norm from a randomized subset from the complete set and 4.72 lower than the rate for all cases where the one or more of the youth's siblings has been arrested.

Table2.9.

	LMean	LCntB	GMean	GCntB
sibarr	1.0	1172	0.2574	1742

Figure 2.11.: Layer 18

Random selection:	Select cases where sibarr = 1
	
Mean: 0.3871	Mean: 0.4274
Std: 0.0133	Std: 0.0083

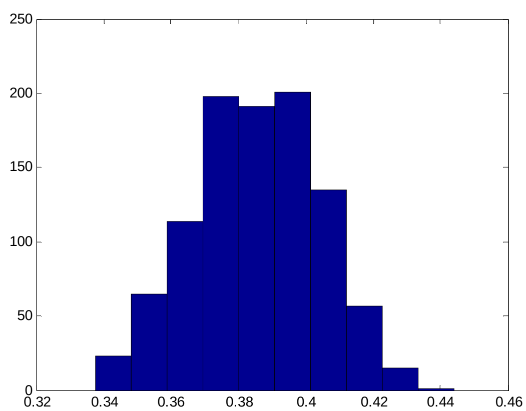
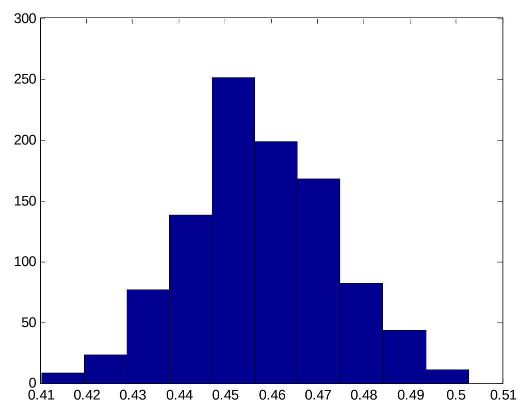
Layer 22

Layer 22 has a recidivism rate of 45.47%, 563 cases and 3 attributes: age of juvenile at instance (age), percent of census block that graduated high school (p_highsch) and youth has previous person offense charges (PriorPersonalChgs). A little over 98% of these youths have had prior personal charges. This layer's recidivism rate is 3.78 standard deviations higher than the norm from a randomized subset from the complete set and 0.2 lower than the rate for all cases where the youth has had prior person offense charges.

Table 2.10.

	LMean	LCntB	GMean	GCntB
age	15.81		15.64	
p_highsch	63.25		58.90	
PriorPersonalChgs	0.9805	552	0.1915	1296

Figure 2.12.: Layer 22

<table>
<tr><td>Random selection:</td><td>Select cases where PriorPersonalChgs = 1</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Mean: 0.3858</td><td>Mean: 0.4578</td></tr>
<tr><td>Std: 0.0185</td><td>Std: 0.0158</td></tr>
</table>

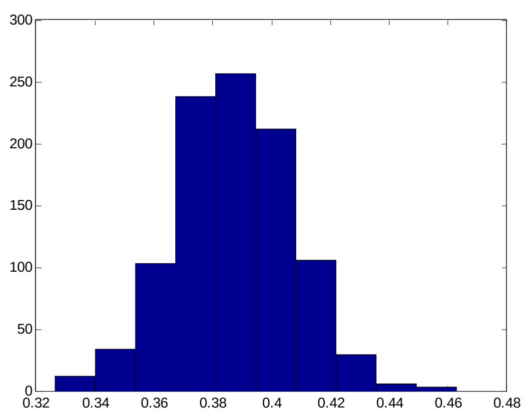
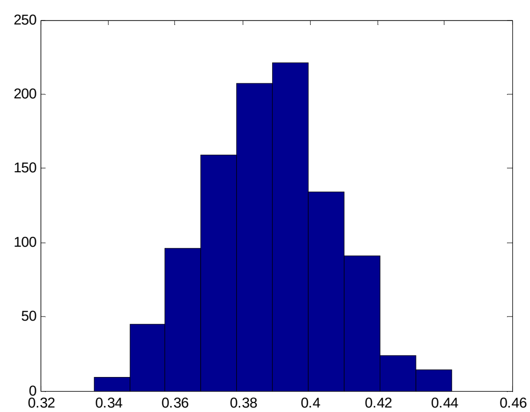
Layer 29

Layer 29 has a recidivism rate of 33.02%, 536 cases and 4 attributes: percent of census block occupants that are African American (p_black), the recidivism density in the youth's census block (gi), number of youths within 1 kilometer of the youth's primary residence at instant (kcnt_1km) and population density near the youth's home (den_person). These are census blocks with much higher than average African American populations and much higher than average population densities. This layer's recidivism rate is 2.83 standard deviations lower than the norm from a randomized subset from the complete set and 3.04 lower than the rate for all cases where the youth's census block is 70% or more African American and the population density is greater than 317 per census block.

Table 2.11.

	LMean	GMean
p_black	85.66	64.66
gi	3.1256	1.6324
kcnt_1km	219.40	185.75
den_person	551.15	426.23

Figure 2.13.: Layer 29

Random selection:	Select cases where p_black >= 70 (mean = 85.66 std = 14.63) AND den_person >= 317 (mean = 551.15 std = 233.83)
	
Mean: 0.3871	Mean: 0.3880
Std: 0.0201	Std: 0.0190

The low recidivism rate of this layer is counterintuitive. Her we have a densely populated set of census blocks, a high percentage of African American residents, which together are associated with an average recidivism rate of 39%. When we add in two spatial data measuring relatively high number s of delinquent peers, the rate drops to 33%.

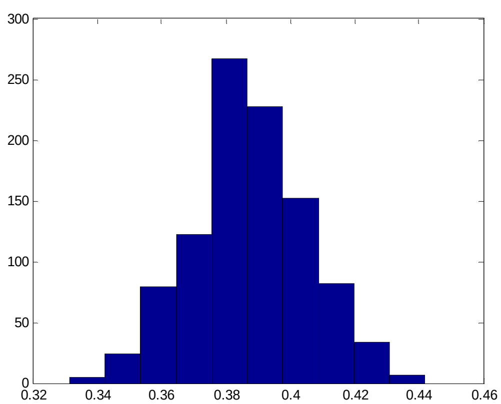
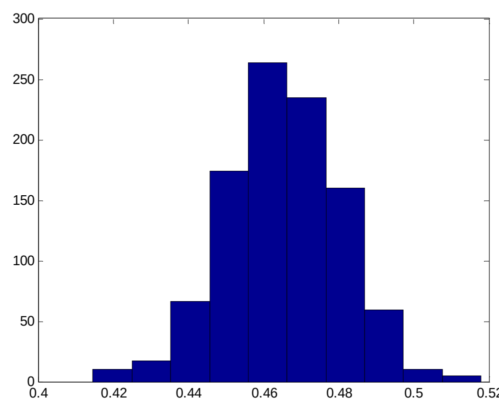
Layer 31

Layer 31 has a recidivism rate of 42.75%, 697 cases and 1 attribute: the youth lived in an institution at instant offense (LiveInstitution). All the youths in this layer were institutionalized at their instant offense. This layer's recidivism rate is 2.23 standard deviations higher than the norm from a randomized subset from the complete set and 2.45 lower than the rate for all cases where the youth lived in an institution at the instance offense.

Table 2.12.

	LMean	LCntB	GMean	GCntB
LiveInstitution	1.00	697	0.3218	2178

Figure 2.14.: Layer 31

Random selection:	Select cases where LiveInstitution = 1
	
Mean: 0.3875	Mean: 0.4654
Std: 0.0179	Std: 0.0155

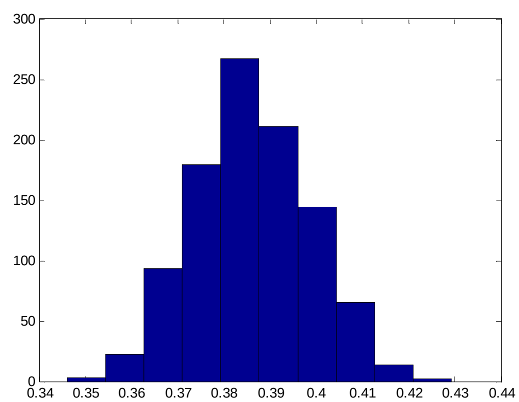
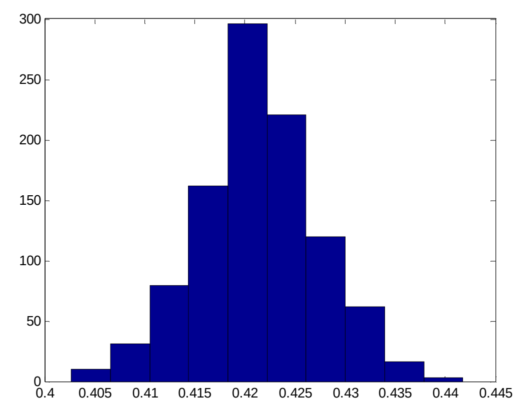
Layer 34

Layer 34 has a recidivism rate of 42.35%, 1202 cases and 2 attributes: youth has had a drug problem (DrugAbuse) and youth has had an alcohol problem (AlcoholAbuse). All the youths in this layer have an alcohol problem and 87% also have a drug problem. This layer's recidivism rate is 2.96 standard deviations higher than the norm from a randomized subset from the complete set and 0.35 lower than the rate for all cases where the youth has alcohol abuse problems.

Table 2.13.

	LMean	LCntB	GMean	GCntB
DrugAbuse	0.8719	1048	0.4634	3136
AlcoholAbuse	1.00	1202	0.2114	1431

Figure 2.15.: Layer 34

<table>
<tr><td>Random selection:</td><td>Select cases where AlcoholAbuse = 1</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>Mean: 0.3865</td><td>Mean: 0.4214</td></tr>
<tr><td>Std: 0.0125</td><td>Std: 0.0060</td></tr>
</table>

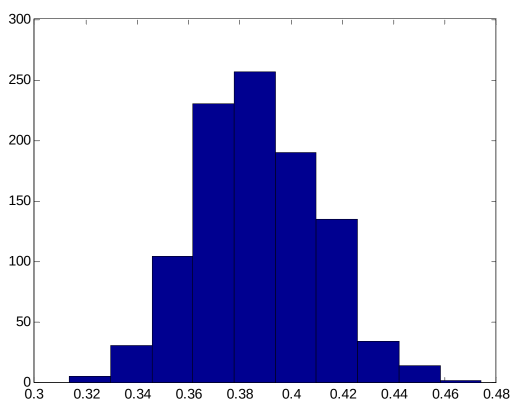
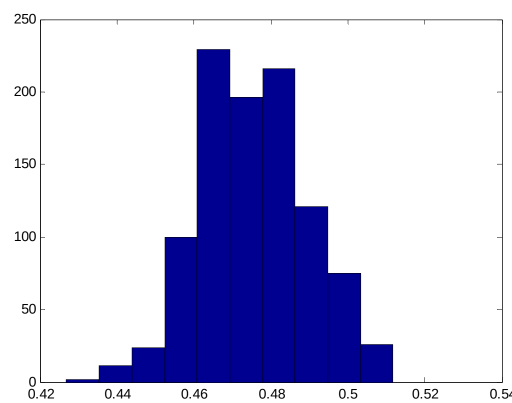
Layer 43

Layer 43 has a recidivism rate of 48.11%, 424 cases and 1 attribute: the youth was on probation at instant offense (Probation). All the youths in this layer were on probation at their instant offense. This layer's recidivism rate is 3.95 standard deviations higher than the norm from a randomized subset from the complete set and 0.94 lower than the rate for all cases where the youth has drug abuse problems. This layer's recidivism rate is 4.94 standard deviations higher than the norm from a randomized subset from the complete set and 0.37 higher than the rate for all cases where the youth was on probation at the instant offense.

Table 2.14.

	LMean	LCntB	GMean	GCntB
Probation	1.00	424	0.0947	641

Figure 2.16.: Layer 43

Random selection:	Select cases where Probation = 1
	
Mean: 0.3875	Mean: 0.4759
Std: 0.0237	Std: 0.0140

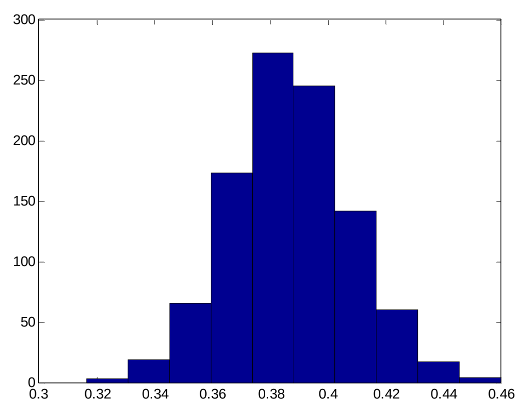
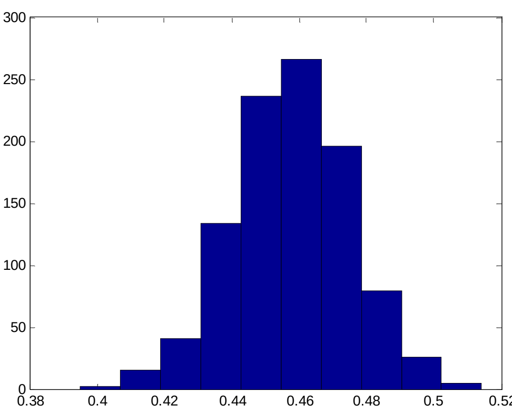
Layer 50

Layer 50 has a recidivism rate of 44.96%, 496 cases and 1 attribute: the youth has previous person offense charges (PriorPersonalChgs). All the youths in this layer were on probation at their instant offense. This layer's recidivism rate is 3.03 standard deviations higher than the norm from a randomized subset from the complete set and 0.46 lower than the rate for all cases where the youth had prior person related charges.

Table 2.15.

	LMean	LCntB	GMean	GCntB
PriorPersonalChgs	1.00	496	0.1915	1296

Figure 2.17.: Layer 50

Random selection:	Select cases where PriorPersonalChgs = 1
	
Mean: 0.3871	Mean: 0.4577
Std: 0.0206	Std: 0.0177

Layer 50 reinforces the layer-recidivism relationship found in Layer 22. Having charges for offenses against persons in the past is related to higher than average recidivism rates.

Layer Comments

Most of these layers have recidivism rates that are significantly different from a same-size unrestricted sample from the complete data set. Many layers also have significantly different recidivism rates when the random selections are restricted by selecting only cases that are within the subgroup bounded by the dominant variable(s). In all layers, some cases were excluded by Plaid by the contributions made by less dominant variables and possibly even the excluded variables. We hope that this has the effect of separating the wheat from the chaff – increasing the signal and reducing the noise.

Attribute Selection

The results from attribute selection are grouped by recidivism type. The tables in this section contain the top 10 predictors based on the Chi-Squared Measure. The variables are sorted in order of significance, with the most significant variable is listed as "Variable 1" and decrease in significance as the variable number increases. The first column lists the data set used for attribute selection. The "All" is the baseline data set that contains all cases. Each record is a set of two rows, where the top row contains the name of the variable and the bottom row contains the value of the Chi-Squared Measure. Gray bars have been added to group each records data to make these tables more readable. The number listed under the data set name is the number of cases within that set.

Ganypet

This variable represents all recidivism offense types, and includes offenses committed between the point of commitment to a specific community program and six months following discharge from that program. The table (2.16) below shows that all layers with a number of cases greater than or equal to 697 have a variable that represents prior educations as one of the top 10 significant variables for predicting recidivism. In fact, prior educations and prior drug arrests are the two most common significant variables in these layers for the prediction of this variable.

For nearly every layer, one of the first five predictors of recidivism in general is drug-related (e.g. prjda=prior drug abuse). For Layer 50, the first predictor is dosage of substance abuse treatment. Only one Layer seems to not have drug-related predictors: Layers 43. These are youths identified in terms of having been on probation at the time they committed their instant offense.

Table 2.16.

Ganypet	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10
ALL	gprjda	drisktot	nincreps	minagepo	scratio	numberdi	Znumberdi	dntot	progcomp	drgarr
6675	156.1068	132.30393	131.21426	123.81337	112.40868	111.81534	111.81534	109.08625	108.51925	107.9734
Layer 1	prjda	nincreps	prjda12	edu_dose	spanish	robbery	Znumberdi	drisktot	suplevio=0	dnp1tot
3357	56.507502	52.919579	50.986613	44.357669	43.944687	43.072745	41.20484	38.333676	35.369963	34.534985
Layer 3	prjda	drgarr	gprjda12	drisktot	numppl=0	suplevio=0	dnp1tot	rpjuv	dntot	DHSAftercare
2279	51.29832	41.15363	40.90835	33.05625	31.2616	29.00537	28.65754	28.1682	27.14417	27.05642
Layer 5	nincreps	scratio	edu_dose	tut_dose	gprjda	drisktot	DHSDay	cms_dose	clientca	dntot
1989	59.99456	56.75651	55.23665	55.09011	53.74559	49.0855	44.20197	43.4619	41.21261	39.53487
Layer 6	progcomp	dntot	prjda	drisktot	nincreps	dnp1tot	sab_dose	arriocd	sexoff	Znumberls
2385	45.66798	45.51488	43.77848	38.97357	38.31563	37.91504	36.38347	36.13804	32.13813	30.76245
Layer 8	schperf	LivesInstitution	numppl=0	drgarr	DoseGrpCoun	gprjda	scratio	DHSAftercare	Zscratio	drisktot
1657	28.89214	25.88801	23.84927	22.40156	22.30557	22.26576	20.35614	20.0711	19.17438	18.85543
Layer 15	drisktot	progcomp	numppl=0	fundtype=1	sibnum	rpjuv	sexoff	jhisda	fcjur	arriocd
641	18.43711	14.9918	13.88207	12.14239	11.13856	10.86594	10.81027	10.72168	10.57048	10.3424
Layer 18	gprjda12	prjda	Probation	drgarr	nincreps	Znumberdi	suplevio=0	LivesInstitution	gnopprob	DHSAftercare
1172	38.7908	36.1459	27.1741	22.5766	22.1027	21.5786	19.3458	19.3321	19.0429	15.8031
Layer 22	nemp=2	edu_dose	resp	InstantPerson	InstantProperty	fcomm	goffc2=2	nemp=0	goffc2=3	schcomit=3625
563	13.9504	12.578	12.5215	12.2241	11.9803	11.2703	11.1267	10.5974	10.3812	8.4915
Layer 29	goffc2=5	fctime	outcome=0	tutoring	gprjda12	outcome=2	commskl	gtinprog	progcomp	nattrec
536	16.95048	16.84034	13.40777	12.58145	11.70894	11.64846	10.96615	10.51306	9.5418	9.13231
Layer 31	nincreps	JuvDrg	drisktot	progcomp	inc_dose	drugoff	dndrugs=1	suplevio=0	gprjda	goffc2=2
697	34.44629	17.02923	16.55703	10.77423	10.40565	10.00393	9.99678	9.57454	9.2155	9.10747
Layer 34	prjda	drisktot	prjda12	suplevio=0	dnedsub2	dntot	minagepo	spanish	age1adj	age1arr
1202	33.1662	23.3603	22.5401	21.2689	20.6464	19.6837	17.2201	16.5796	15.8137	15.4532
Layer 43	clientca	DoseGrpCoun	program=120	minagepo	v_dang	grc_dose	fundamt	avglos	propoff	autothft
424	23.86458	19.6238	15.53614	15.48617	15.0557	14.06625	13.88145	13.82233	10.9703	10.68253
Layer 50	DoseSubAbuse	avglos	marstatp=2	r_lowiq	v_noncom	dae_dose	sibnum	goohpnum	numppl=0	adjcts
496	14.5224	12.8009	12.0782	11.0571	9.0824	8.5142	8.5142	8.3227	8.2021	8.1607

XDrugs

This variable represents the drug recidivism type. In a case of drug recidivism, all data sets have significant indicators that represent a history of drug arrests. What's most interesting about these results are that the partitions, created by the layers, appear to have caused a great increase in the value of drug related histories with respect to drug recidivism. Notice that the first data set, “ALL”, has the variable “drgarr”, which is the number of juvenile arrests for a drug offense, is listed as the ninth most significant variable. In 11 out of the 13 layers, a drug related variable is the number one most significant predictor for drug related recidivism. Five layers have “JuvDrg”, any juvenile drug arrests, five more have “drgarr”, and one has “InstantDrug” as the most significant predictor. In fact, many layers have multiple occurrences of drug related predictors for drug recidivism.

The predictors in these layers strongly suggest that drug recidivists have specialized in drug offending. Layer 43 is particularly interesting because of its ethnic attributes. Being Hispanic, or living in a Hispanic neighborhood is strongly associated with drug offending when recidivating. Again, these are youths who were on probation at the time of their instant offense.

Table 2.17.

Xdrugs	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10
ALL	gprjda	drisktot	nincreps	minagepo	scratio	numberdi	dntot	progcomp	drgarr	sab_dose
6675	156.1068	132.30393	131.21426	123.81337	112.40868	111.81534	109.08625	108.51925	107.9734	102.6354
Layer 1	JuvDrg	InstantDrug	goffc2=1	age	InstantPerson	age1arr	DoseSubAbuse	jhisda	totcharg	dnemp=2
3357	89.6417	79.2225	46.6903	35.6931	32.1182	25.7493	23.2982	23.0975	23.0662	22.8927
Layer 3	drgarr	InstantDrug	spanish	goffc2=1	pct_hisp	Hispanic	fundamt	DHSAftercare	livarr=1	schperf
2279	144.8491	61.891	46.5363	46.0563	39.1037	37.0168	33.483	33.0122	32.6302	30.688
Layer 5	drgarr	arriocd	rpjuv	numppl=0	scratio	drisktot	DoseSubAbuse	Zscratio	outcome=0	altpclas=4
1989	77.0047	28.27022	26.5308	25.3609	22.42054	22.07234	21.62423	21.55336	20.39167	19.81561
Layer 6	JuvDrg	sab_dose	Znumberdi	age	rpjuv	DoseSubAbuse	robbery	numppl=0	weapons	spanish
2385	63.61471	37.61612	33.48029	32.46851	31.28179	27.40006	26.65617	25.77154	24.78454	22.13403
Layer 8	drgarr	DHSAftercare	scratio	altpclas=4	schperf	dnvotec	LivesInstitution	livarr=1	drisktot	rpjuv
1657	37.1601	20.9565	20.2247	19.0969	18.6196	17.387	17.2781	16.9508	15.4978	15.4682
Layer 15	InstantDrug	goffc2=1	drgarr	ndiff	fintcom	totcharg	jhisda	program=57	disnatt	esteem_2=1
641	49.81003	41.43674	35.85942	23.73644	15.77466	14.1951	14.00649	12.3828	12.17208	11.97025
Layer 18	drgarr	goffc2=1	InstantDrug	pct_hisp	Hispanic	spanish	dndrugs=0	gprjda12	recprop	dnemp=2
1172	75.217494	46.517822	44.300799	33.639381	22.658944	20.464579	20.0006	17.346618	14.919479	14.443135
Layer 22	JuvDrg	InstantDrug	goffc2=1	InstantPerson	resp	program2=131	goffc2=2	csrcinc4	ndisrpt	OutofHome
563	29.00886	18.16336	16.33555	14.11829	12.76083	11.48696	11.04281	8.54644	8.36949	8.26025
Layer 29	ictime	drisktot	dntot	age	dnp1tot	minagepo	fctime	goffc2=1	nsasub	disptype=4
536	58.6584	27.6874	17.6056	15.0998	14.8704	13.1242	13.0271	12.9753	12.2948	11.7542
Layer 31	JuvDrg	InstantDrug	goffc2=1	arriocd	numppl=4	edstdisp=0	progcomp	dnp1tot	v_drugsl	DoseSubAbuse
697	44.7991	36.1036	34.0196	15.7039	13.562	10.1758	9.3849	9.1658	9.0737	8.5606
Layer 34	drgarr	pct_hisp	InstantDrug	spanish	goffc2=1	prjda12	numcount	Hispanic	prjda	schperf
1202	56.4787	29.2024	28.8135	26.4597	22.9357	22.6214	17.0437	13.7358	12.7914	12.4993
Layer 43	pct_hisp	drgarr	Hispanic	program2=57	goffc2=1	InstantDrug	rpjuv	schcomit=3625	minagepo	stabfam
424	26.8946	26.7221	17.7699	15.5791	14.8569	13.9168	11.3754	10.3077	7.6697	7.5046
Layer 50	JuvDrg	DoseSubAbuse	avglos	hpabusef=0	nemp=2	nemp=0	rfmviol4	DHSAftercare	gedip	reinsch
496	18.1639	17.2914	12.7856	11.2437	10.325	10.3003	10.2955	9.5087	8.9605	8.8292

XPerson

This variable represents person related recidivism. The most frequently occurring variables for this type of recidivism are instant or prior person crimes (PriorPerson, pcspo, gpcspo, goffc2=2), and a history of out of home placement (numppl, age1oop, OutOfHome). These attributes suggest family mismanagement and stressful environments for these youths.

Table 2.18.

XPerson	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10
ALL	tinprog	progtime	gtinprog	age1arr	nincreps	age1oop	PriorPerson	pcspo	gpcspo	numppl=6
6675	35.36858	34.14113	33.14747	30.35394	29.87397	28.77383	25.10006	25.10006	25.10006	24.61218
Layer 1	numppl=6	age1arr	jhismh	PriorPerson	pcspo	gtinprog	age1adj	program=781	schcomit=4125	par_dose
3357	26.667	21.1003	17.1565	11.7062	11.7062	11.5696	11.1133	10.4573	10.1466	9.1368
Layer 3	numcount	JuvDrg	InstantPerson	Black	r_xviol	nhilfunc	goffc2=2	program2=185	program2=165	schcomit=125
2279	20.5082	14.5889	14.2596	12.9017	12.5474	11.9967	11.1898	10.2496	10.2496	10.2496
Layer 5	pcspo	age1oop	v_fabide	numppl=6	icna	v_arson	nhilfunc	fc_dose	wrk_dose	resp
1989	19.2897	18.8695	18.7738	17.4167	15.7392	13.3851	12.8971	12.8258	12.8258	12.7283
Layer 6	program2=185	schcomit=17143	arriocd	PriorPerson	gpcspo	wnselfcn	gprjda	r_meds	dnpersub	peropp=1
2385	15.8067	15.3439	15.3026	14.46	14.46	13.071	10.8585	10.7236	10.0753	9.8581
Layer 8	program=73	program2=126	drisktot	schcomit=4125	dntot	disnatt	wnresist	conarr=45	typdisch=3	dnemstab=2
1657	12.5531	12.5531	9.9375	7.3992	7.0573	6.4843	6.3586	5.5916	5.363	4.7928
Layer 15	prenum	numcount	behavpr	arson	terrthr	OutofHome	v_lattnd	simpass	progcomp	schcomit=22857
641	16.5736	16.5736	11.0454	10.0966	9.8153	9.8051	9.2152	9.0627	8.3259	8.2738
Layer 18	navsex	PriorPerson	gpcspo	dana	gschinv=4	schcomit=1	comminv=5	schcomit=4125	schinv=1	gschinv=1
1172	16.7181	10.0592	10.0592	8.4581	8.4144	8.142	7.8412	7.8412	7.6166	7.6166
Layer 22	gohptime	race=4	JuvDrg	InstantDrug	drugoff	gtcharge	nhilfunc	numppl=5	program2=78	program2=156
563	15.1314	14.7094	10.5922	8.6481	7.7234	7.4269	7.3415	7.3415	7.3415	7.3415
Layer 29	schinv=2	schcomit=4125	schcomit=15	disptype=0	commres	famneeds	mandarin	edstdisp=2	ethiop	drugoff
536	14.6884	9.8986	9.253	9.253	6.603	6.1926	5.6646	5.6646	5.4016	5.044
Layer 31	fctime	prenum	truancy	gtinprog	r_homi	gcna	outcome=0	program2=78	mined=2	truexpl
697	30.301	18.524	11.818	11.341	11.166	10.31	9.249	9.249	9.249	9.249
Layer 34	numppl=7	policedist	age1arr	numppl=6	sibarr	pcspo	gpcspo	resident	edstdisp=2	mhinoff
1202	19.2645	17.7895	16.4453	15.0554	12.2546	11.3889	11.3889	9.1898	8.591	7.9793
Layer 43	altpclas=3	pev_dose	fundtype=3	DHSCounseling	mined=4	grt_dose	program=1	sibarr	v_dang	nloner
424	19.2954	13.0314	13.0314	13.0314	13.0314	12.8329	12.8329	12.773	6.7152	6.4538
Layer 50	program=781	arson	schcomit=38333	program=165	altpcls2=7	program=135	livarr=0	smotdec	sfatdec	program2=78
496	10.1633	7.4527	6.2333	6.2333	6.2333	6.2333	6.2333	6.2333	6.2333	6.2333

XProperty

This variable represents property related recidivism. The baseline set and six of the layers contain variables that represent that the instant offense was property related (InstantProperty, goffc2=3). The predictors for this type of reoffending are, however, more diverse, indicating a less clear pattern for this category of offense.

Table 2.19.

Xproperty	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10
ALL	InstantProperty	goffc2=3	oohptime	nincreps	progcomp	drisktot	gprjda	dntot	ndiff	InstantDrug
6675	56.985	55.3807	50.8501	44.3407	39.8211	36.5905	28.4371	26.7467	26.6711	26.2187
Layer 1	progcomp	goffc2=3	pct_vacant	InstantProperty	prjda12	prjda	dnp1tot	sab_dose	edu_dose	Znumberls
3357	37.2254	32.8505	28.6709	27.2227	24.3971	19.4943	18.1943	18.1505	17.7723	17.5719
Layer 3	InstantDrug	aftschoold	InstantProperty	ndiff	JuvDrg	schcomit=15	drisktot	age1adj	gprjda	prjda
2279	21.6386	21.3411	19.6652	18.2428	18.0857	17.5724	15.9386	15.9238	13.3873	13.3873
Layer 5	nincreps	progcomp	sab_dose	edu_dose	fundamt	DHSDay	ftdisty=5	prjda12	DoseGrpCoun	numberpr
1989	32.6171	15.7446	14.4054	14.0461	13.0151	13.0101	12.806	12.5703	11.5027	10.8016
Layer 6	drisktot	dntot	progcomp	dnp1tot	prjda	age1oop	fc_dose	wrk_dose	goohpnum	dnedsub2
2385	20.26356	17.25238	16.87167	15.71898	15.03331	14.62964	14.03403	14.03403	13.51583	13.25212
Layer 8	program2=771	sfatdec	sed_dose	curfviol	numppl=6	age1arr	suplevio=0	sattach=45	weapinv	gcomminv=3
1657	22.8612	13.9436	11.4236	11.4236	11.4236	10.7148	7.3043	7.2003	7.1362	6.4943
Layer 15	arriocd	dnpersub	comminv=1	gcomminv=1	outcome=2	ftdisty=1	fcjur	gcomminv=3	schcomit=44286	program2=782
641	10.3487	9.53	9.4784	9.4784	8.7358	8.4502	7.8255	7.0424	6.8997	6.8997
Layer 18	race=3	dnintab=2	Zscratio	Znumberls	InstantProperty	ftdisty=5	prjda	goffc2=3	program=57	weapons
1172	16.0723	15.7116	15.3535	15.3535	13.5118	13.0209	12.4116	12.3102	10.6808	10.1021
Layer 22	InstantProperty	goffc2=3	schcomit=225	sattach=1	comphsip	IQ=3	DHSSIL	altpclas=11	program=89	Modality=6
563	20.8424	19.0938	10.9988	10.1322	9.7148	7.3892	7.3415	7.3415	7.3415	7.3415
Layer 29	nsasub	nsdev	ind_dose	job_dose	edstdisp=2	nprfam=0	goffc2=3	prjda12	ftdisty=0	nattrec
536	17.0163	15.58464	14.99306	14.48538	13.02576	12.73095	12.04059	11.75074	11.1553	10.5219
Layer 31	dnvotec	schcomit=15	InstantProperty	goffc2=3	ftdisty=7	subabu	age1arr	conarr=45	faminc=1	schcomit=45
697	9.5852	9.2486	9.2254	8.4081	7.2661	6.4709	5.5824	3.4555	3.0416	2.4599
Layer 34	ndiff	dnldis=2	suplevio=4	gcomminv=2	schinv=5	dnvotec	progcomp	dnedsub2	dnemstab=2	wntooimm
1202	21.1039	13.9031	11.6564	10.9341	10.7932	10.665	10.233	9.1869	9.1292	8.9482
Layer 43	schcomit=34286	gprjda	csrcinc2	program=178	schcomit=46667	disptype=7	numppl=5	gtinprog	gschinv=3	schinv=1
424	13.9531	9.5886	7.617	6.9599	6.9599	6.9599	6.9599	6.1334	5.6433	5.4667
Layer 50	ad_person	nhlthsub	schcomit=46667	program2=782	program2=77	famhouse=6	ftdisty=0	comphsip	hpabusef=2	decsk
496	18.136	14.992	7.481	7.481	7.481	5.964	4.294	2.793	2.359	2.359

Neural Network Analysis

The next question we address is whether the co-clusters or layers improve our ability to predict recidivism. Our neural network analysis is comprised of 112 trials, 56 for prediction and 56 for modeling, and is shown in a series of 8 tables. There are 4 for prediction and 4 for modeling. Each method has a table for each of the 4 target variables. As with attribute selection, rows are the layers and the columns are types of information about the layers. The first column identifies the data set. Columns 2 through 5 are the confusion matrix [Kohavi] for the experiment on the given rows layer (TP is true positive, FP is false positive, FN is false negative and TN is true negative). Columns 6, 7 and 8 are the number of negatives, positives and total cases within the layer. The 9th column is the recidivism ratio and the 10th column is the overall accuracy. The 11th, 12th and 13th columns are the sensitivity (true positives divided by all positives), specificity (true negatives divided by all negatives) and accuracy based on sensitivity and specificity ((sensitivity + specificity)/2). The last two columns represent the deviation from the baseline accuracies in the row for the "All" data set for each layer.

Prediction

The table below shows the predictions for the variable that represents all recidivism types. The results for all recidivism types are typically less for the partitioned layers then for the baseline. On average the partitioned models are 5.68% less accurate and 2.04% less accurate using the specificity and sensitivity measure. Most of the partitioned models are less accurate than the baseline for both types of accuracy.

Table 2.20.

Predictive Accuracy All Recidivism (ganypet)														
	TN	FP	FN	TP	Negative	Positive	Total	RecidRatio	Accuracy	Sensitivity	Specificity	Saccuracy	AccDiff	SaccDif
All Data	673	122	393	147	795	540	1335	40.45%	61.42%	27.22%	84.65%	55.94%		
Layer 01	360	84	158	62	444	220	664	33.13%	63.55%	28.18%	81.08%	54.63%	2.13%	-1.31%
Layer 03	143	112	93	102	255	195	450	43.33%	54.44%	52.31%	56.08%	54.19%	-6.98%	-1.75%
Layer 05	136	98	71	89	234	160	394	40.61%	57.11%	55.63%	58.12%	56.87%	-4.32%	0.93%
Layer 06	251	62	106	52	313	158	471	33.55%	64.33%	32.91%	80.19%	56.55%	2.91%	0.61%
Layer 08	83	90	71	84	173	155	328	47.26%	50.91%	54.19%	47.98%	51.09%	-10.51%	-4.85%
Layer 15	46	29	36	17	75	53	128	41.41%	49.22%	32.08%	61.33%	46.70%	-12.20%	-9.23%
Layer 18	75	72	45	38	147	83	230	36.09%	49.13%	45.78%	51.02%	48.40%	-12.29%	-7.54%
Layer 22	40	17	30	25	57	55	112	49.11%	58.04%	45.45%	70.18%	57.81%	-3.39%	1.88%
Layer 29	45	25	18	16	70	34	104	32.69%	58.65%	47.06%	64.29%	55.67%	-2.77%	-0.27%
Layer 31	48	34	25	32	82	57	139	41.01%	57.55%	56.14%	58.54%	57.34%	-3.87%	1.40%
Layer 34	81	53	55	49	134	104	238	43.70%	54.62%	47.12%	60.45%	53.78%	-6.80%	-2.16%
Layer 43	22	21	15	26	43	41	84	48.81%	57.14%	63.41%	51.16%	57.29%	-4.28%	1.35%
Layer 50	28	19	30	21	47	51	98	52.04%	50.00%	41.18%	59.57%	50.38%	-11.42%	-5.56%
Average	104.4615	55.07692	57.92308	47.15385	159.5385	105.0769	264.6154	39.71%	55.75%	46.26%	61.54%	53.90%	-5.68%	-2.04%

The next table shows the predictive results for drug related recidivism. Notice that the average accuracies are slightly higher for this type of recidivism than for all recidivism. More than half, 8/13 trials, have higher overall accuracy and 10/13 trials have higher accuracy based on sensitivity and specificity.

Table 2.21.

Predictive Accuracy Drug Recidivism (XDrugs)

	TN	FP	FN	TP	Negative	Positive	Total	RecidRatio	Accuracy	Sensitivity	Specificity	Saccuracy	AccDiff	SaccDif
All Data	1024	115	169	27	1139	196	1335	14.68%	78.73%	13.78%	89.90%	51.84%		
Layer 01	550	50	59	5	600	64	664	9.64%	83.58%	7.81%	91.67%	49.74%	4.86%	-2.10%
Layer 03	312	56	59	23	368	82	450	18.22%	74.44%	28.05%	84.78%	56.42%	-4.28%	4.58%
Layer 05	335	23	32	4	358	36	394	9.14%	86.04%	11.11%	93.58%	52.34%	7.31%	0.50%
Layer 06	416	27	23	5	443	28	471	5.94%	89.38%	17.86%	93.91%	55.88%	10.66%	4.04%
Layer 08	195	50	53	30	245	83	328	25.30%	68.60%	36.14%	79.59%	57.87%	-10.13%	6.03%
Layer 15	103	8	13	4	111	17	128	13.28%	83.59%	23.53%	92.79%	58.16%	4.87%	6.32%
Layer 18	180	15	30	5	195	35	230	15.22%	80.43%	14.29%	92.31%	53.30%	1.71%	1.46%
Layer 22	93	11	8	0	104	8	112	7.14%	83.04%	0.00%	89.42%	44.71%	4.31%	-7.13%
Layer 29	93	5	4	2	98	6	104	5.77%	91.35%	33.33%	94.90%	64.12%	12.62%	12.28%
Layer 31	93	20	16	10	113	26	139	18.71%	74.10%	38.46%	82.30%	60.38%	-4.63%	8.54%
Layer 34	175	24	27	12	199	39	238	16.39%	78.57%	30.77%	87.94%	59.35%	-0.16%	7.51%
Layer 43	62	7	12	3	69	15	84	17.86%	77.38%	20.00%	89.86%	54.93%	-1.35%	3.09%
Layer 50	84	5	9	0	89	9	98	9.18%	85.71%	0.00%	94.38%	47.19%	6.99%	-4.65%
Average	207	23.15385	26.53846	7.923077	230.1538	34.46154	264.6154	13.02%	81.25%	20.10%	89.80%	54.95%	2.52%	3.11%

Person related recidivism prediction did not perform as well as drug. Partitioning the data did improve performance over all recidivism, however: more than half of the layers outperformed the baseline accuracy for both methods of comparison.

Table 2.22.

Predictive Accuracy Person Recidivism (XPerson)

	TN	FP	FN	TP	Negative	Positive	Total	RecidRatio	Accuracy	Sensitivity	Specificity	Saccuracy	AccDiff	SaccDif
All Data	1154	52	119	10	1206	129	1335	9.66%	87.19%	7.75%	95.69%	51.72%		
Layer 01	569	25	66	4	594	70	664	10.54%	86.30%	5.71%	95.79%	50.75%	-0.90%	-0.97%
Layer 03	388	18	37	7	406	44	450	9.78%	87.78%	15.91%	95.57%	55.74%	0.59%	4.02%
Layer 05	320	17	52	5	337	57	394	14.47%	82.49%	8.77%	94.96%	51.86%	-4.70%	0.14%
Layer 06	401	10	54	6	411	60	471	12.74%	86.41%	10.00%	97.57%	53.78%	-0.78%	2.06%
Layer 08	301	2	25	0	303	25	328	7.62%	91.77%	0.00%	99.34%	49.67%	4.58%	-2.05%
Layer 15	116	0	11	1	116	12	128	9.38%	91.41%	8.33%	100.00%	54.17%	4.22%	2.45%
Layer 18	202	7	19	2	209	21	230	9.13%	88.70%	9.52%	96.65%	53.09%	1.50%	1.37%
Layer 22	97	0	14	1	97	15	112	13.39%	87.50%	6.67%	100.00%	53.33%	0.31%	1.61%
Layer 29	91	4	9	0	95	9	104	8.65%	87.50%	0.00%	95.79%	47.89%	0.31%	-3.83%
Layer 31	116	9	14	0	125	14	139	10.07%	83.45%	0.00%	92.80%	46.40%	-3.74%	-5.32%
Layer 34	200	17	18	3	217	21	238	8.82%	85.29%	14.29%	92.17%	53.23%	-1.90%	1.51%
Layer 43	70	0	14	0	70	14	84	16.67%	83.33%	0.00%	100.00%	50.00%	-3.86%	-1.72%
Layer 50	78	0	20	0	78	20	98	20.41%	79.59%	0.00%	100.00%	50.00%	-7.60%	-1.72%
Average	226.8462	8.384615	27.15385	2.230769	235.2308	29.38462	264.6154	11.10%	86.27%	6.09%	96.97%	51.53%	-0.92%	-0.19%

The averages for the predictions for property recidivism are higher than all recidivism and many layers outperformed the baseline for both comparative methods.

Table 2.23.

Predictive Accuracy Property Recidivism (XProperty)														
	TN	FP	FN	TP	Negative	Positive	Total	RecidRatio	Accuracy	Sensitivity	Specificity	Saccuracy	AccDiff	SaccDif
All Data	1112	68	146	9	1180	155	1335	11.61%	83.97%	5.81%	94.24%	50.02%		
Layer 01	558	53	45	8	611	53	664	7.98%	85.24%	15.09%	91.33%	53.21%	1.27%	3.19%
Layer 03	367	31	46	6	398	52	450	11.56%	82.89%	11.54%	92.21%	51.87%	-1.08%	1.85%
Layer 05	309	33	47	5	342	52	394	13.20%	79.70%	9.62%	90.35%	49.98%	-4.27%	-0.04%
Layer 06	390	37	43	1	427	44	471	9.34%	83.01%	2.27%	91.33%	46.80%	-0.96%	-3.22%
Layer 08	302	2	23	1	304	24	328	7.32%	92.38%	4.17%	99.34%	51.75%	8.41%	1.73%
Layer 15	100	10	16	2	110	18	128	14.06%	79.69%	11.11%	90.91%	51.01%	-4.28%	0.99%
Layer 18	193	18	18	1	211	19	230	8.26%	84.35%	5.26%	91.47%	48.37%	0.38%	-1.66%
Layer 22	92	0	19	1	92	20	112	17.86%	83.04%	5.00%	100.00%	52.50%	-0.93%	2.48%
Layer 29	87	3	11	3	90	14	104	13.46%	86.54%	21.43%	96.67%	59.05%	2.57%	9.03%
Layer 31	124	4	11	0	128	11	139	7.91%	89.21%	0.00%	96.88%	48.44%	5.24%	-1.58%
Layer 34	197	15	23	3	212	26	238	10.92%	84.03%	11.54%	92.92%	52.23%	0.06%	2.21%
Layer 43	76	1	7	0	77	7	84	8.33%	90.48%	0.00%	98.70%	49.35%	6.51%	-0.67%
Layer 50	84	0	13	1	84	14	98	14.29%	86.73%	7.14%	100.00%	53.57%	2.76%	3.55%
Average	221.4615	15.92308	24.76923	2.461538	237.3846	27.23077	264.6154	10.29%	85.18%	8.01%	94.78%	51.40%	1.21%	1.37%

Modeling

The modeling for all recidivism shows that the layers consistently outperformed the baseline accuracy with an average overall accuracy improvement of 14.18% and an 18.19% improvement when considering sensitivity and specificity. Thus, prediction of overall recidivism is improved by clustering cases empirically.

Table 2.24.

Model Accuracy All Recidivism (ganypet)														
	TN	FP	FN	TP	Negative	Positive	Total	RecidRatio	Accuracy	Sensitivity	Specificity	Saccuracy	AccDiff	SaccDif
All Data	3776	315	1529	1055	4091	2584	6675	38.71%	72.37%	40.83%	92.30%	66.56%		
Layer 01	2005	144	516	655	2149	1171	3320	35.27%	80.12%	55.94%	93.30%	74.62%	7.75%	8.05%
Layer 03	1246	44	312	647	1290	959	2249	42.64%	84.17%	67.47%	96.59%	82.03%	11.80%	15.46%
Layer 05	1071	111	284	503	1182	787	1969	39.97%	79.94%	63.91%	90.61%	77.26%	7.56%	10.70%
Layer 06	1541	28	305	483	1569	788	2357	33.43%	85.87%	61.29%	98.22%	79.75%	13.50%	13.19%
Layer 08	836	49	201	553	885	754	1639	46.00%	84.75%	73.34%	94.46%	83.90%	12.37%	17.34%
Layer 15	372	6	7	254	378	261	639	40.85%	97.97%	97.32%	98.41%	97.87%	25.59%	31.30%
Layer 18	701	4	74	372	705	446	1151	38.75%	93.22%	83.41%	99.43%	91.42%	20.85%	24.86%
Layer 22	210	95	32	221	305	253	558	45.34%	77.24%	87.35%	68.85%	78.10%	4.87%	11.54%
Layer 29	333	18	53	118	351	171	522	32.76%	86.40%	69.01%	94.87%	81.94%	14.02%	15.37%
Layer 31	393	6	12	285	399	297	696	42.67%	97.41%	95.96%	98.50%	97.23%	25.04%	30.66%
Layer 34	667	19	42	461	686	503	1189	42.30%	94.87%	91.65%	97.23%	94.44%	22.50%	27.88%
Layer 43	196	23	59	143	219	202	421	47.98%	80.52%	70.79%	89.50%	80.14%	8.15%	13.58%
Layer 50	217	55	30	189	272	219	491	44.60%	82.69%	86.30%	79.78%	83.04%	10.31%	16.48%
Average	752.9231	46.30769	148.2308	375.6923	799.2308	523.9231	1323.154	39.60%	86.55%	77.21%	92.29%	84.75%	14.18%	18.19%

Drug recidivism modeling also shows improvement when compared to the baseline(Table 2.25) . Overall accuracy is higher in most layers but only 8/13 trials fared better for the sensitivity and specificity base accuracy measure.

Table 2.25.

Model Accuracy Drug Recidivism (XDrugs)														
	TN	FP	FN	TP	Negative	Positive	Total	RecidRatio	Accuracy	Sensitivity	Specificity	Saccuracy	AccDiff	SaccDif
All Data	5697	87	428	463	5784	891	6675	13.35%	92.28%	51.96%	98.50%	75.23%		
Layer 01	2965	6	83	266	2971	349	3320	10.51%	97.32%	76.22%	99.80%	88.01%	5.03%	12.78%
Layer 03	1824	10	119	296	1834	415	2249	18.45%	94.26%	71.33%	99.45%	85.39%	1.98%	-2.62%
Layer 05	1771	3	51	144	1774	195	1969	9.90%	97.26%	73.85%	99.83%	86.84%	4.97%	1.45%
Layer 06	2158	0	51	148	2158	199	2357	8.44%	97.84%	74.37%	100.00%	87.19%	5.55%	0.35%
Layer 08	1188	41	98	312	1229	410	1639	25.02%	91.52%	76.10%	96.66%	86.38%	-0.77%	-0.81%
Layer 15	567	0	6	66	567	72	639	11.27%	99.06%	91.67%	100.00%	95.83%	6.78%	9.45%
Layer 18	991	0	16	144	991	160	1151	13.90%	98.61%	90.00%	100.00%	95.00%	6.33%	-0.83%
Layer 22	465	10	8	75	475	83	558	14.87%	96.77%	90.36%	97.89%	94.13%	4.49%	-0.87%
Layer 29	481	0	7	34	481	41	522	7.85%	98.66%	82.93%	100.00%	91.46%	6.37%	-2.66%
Layer 31	528	42	30	96	570	126	696	18.10%	89.66%	76.19%	92.63%	84.41%	-2.63%	-7.05%
Layer 34	1017	1	20	151	1018	171	1189	14.38%	98.23%	88.30%	99.90%	94.10%	5.95%	9.69%
Layer 43	354	2	12	53	356	65	421	15.44%	96.67%	81.54%	99.44%	90.49%	4.39%	-3.61%
Layer 50	429	0	20	42	429	62	491	12.63%	95.93%	67.74%	100.00%	83.87%	3.64%	-6.62%
Average	1133.692	8.846154	40.07692	140.5385	1142.538	180.6154	1323.154	13.65%	96.29%	80.05%	98.89%	89.47%	4.01%	0.66%

Predictions for person recidivism offenses (Table 2.26) were not improved by clustering cases: less than half (7/13 trials) of the layers outperformed the baseline, and the average improved accuracy was -7.24%. The average improved accuracy for property offenders (Table 2.27), however, are higher than was found for person offenders, with improved accuracy ranging as high as 23.26 percent (Layer 29).

Table 2.26

Model Accuracy Person Recidivism (XPerson)														
	TN	FP	FN	TP	Negative	Positive	Total	RecidRatio	Accuracy	Sensitivity	Specificity	Saccuracy	AccDiff	SaccDif
All Data	5995	19	323	338	6014	661	6675	9.90%	94.88%	51.13%	99.68%	75.41%		
Layer 01	2975	9	154	182	2984	336	3320	10.12%	95.09%	54.17%	99.70%	76.93%	0.21%	1.52%
Layer 03	2007	42	93	107	2049	200	2249	8.89%	94.00%	53.50%	97.95%	75.73%	-0.88%	0.32%
Layer 05	1756	10	83	120	1766	203	1969	10.31%	95.28%	59.11%	99.43%	79.27%	0.40%	3.86%
Layer 06	2054	31	180	92	2085	272	2357	11.54%	91.05%	33.82%	98.51%	66.17%	-3.83%	-9.24%
Layer 08	1502	16	93	28	1518	121	1639	7.38%	93.35%	23.14%	98.95%	61.04%	-1.53%	-14.37%
Layer 15	557	13	46	23	570	69	639	10.80%	90.77%	33.33%	97.72%	65.53%	-4.11%	-9.88%
Layer 18	1016	9	92	34	1025	126	1151	10.95%	91.23%	26.98%	99.12%	63.05%	-3.65%	-12.36%
Layer 22	490	1	56	11	491	67	558	12.01%	89.78%	16.42%	99.80%	58.11%	-5.09%	-17.30%
Layer 29	469	2	41	10	471	51	522	9.77%	91.76%	19.61%	99.58%	59.59%	-3.11%	-15.82%
Layer 31	627	1	28	40	628	68	696	9.77%	95.83%	58.82%	99.84%	79.33%	0.96%	3.92%
Layer 34	1076	1	33	79	1077	112	1189	9.42%	97.14%	70.54%	99.91%	85.22%	2.26%	9.81%
Layer 43	364	0	46	11	364	57	421	13.54%	89.07%	19.30%	100.00%	59.65%	-5.80%	-15.76%
Layer 50	423	0	59	9	423	68	491	13.85%	87.98%	13.24%	100.00%	56.62%	-6.89%	-18.79%
Average	1178.154	10.38462	77.23077	57.38462	1188.538	134.6154	1323.154	10.17%	92.49%	37.08%	99.27%	68.17%	-2.39%	-7.24%

Table 2.27.

Model Accuracy Property Recidivism (XProperty)														
	TN	FP	FN	TP	Negative	Positive	Total	RecidRatio	Accuracy	Sensitivity	Specificity	Saccuracy	AccDiff	SaccDif
All Data	5931	26	419	299	5957	718	6675	10.76%	93.33%	41.64%	99.56%	70.60%		
Layer 01	2965	10	100	245	2975	345	3320	10.39%	96.69%	71.01%	99.66%	85.34%	3.35%	14.74%
Layer 03	2011	0	61	177	2011	238	2249	10.58%	97.29%	74.37%	100.00%	87.18%	3.95%	16.58%
Layer 05	1641	33	187	108	1674	295	1969	14.98%	88.83%	36.61%	98.03%	67.32%	-4.51%	-3.28%
Layer 06	2134	4	55	164	2138	219	2357	9.29%	97.50%	74.89%	99.81%	87.35%	4.16%	16.75%
Layer 08	1504	3	117	15	1507	132	1639	8.05%	92.68%	11.36%	99.80%	55.58%	-0.65%	-15.02%
Layer 15	555	3	33	48	558	81	639	12.68%	94.37%	59.26%	99.46%	79.36%	1.03%	8.76%
Layer 18	1034	6	36	75	1040	111	1151	9.64%	96.35%	67.57%	99.42%	83.50%	3.02%	12.89%
Layer 22	484	7	42	25	491	67	558	12.01%	91.22%	37.31%	98.57%	67.94%	-2.11%	-2.66%
Layer 29	465	0	7	50	465	57	522	10.92%	98.66%	87.72%	100.00%	93.86%	5.33%	23.26%
Layer 31	628	0	63	5	628	68	696	9.77%	90.95%	7.35%	100.00%	53.68%	-2.39%	-16.93%
Layer 34	1045	2	52	90	1047	142	1189	11.94%	95.46%	63.38%	99.81%	81.59%	2.13%	10.99%
Layer 43	362	6	34	19	368	53	421	12.59%	90.50%	35.85%	98.37%	67.11%	-2.83%	-3.49%
Layer 50	433	0	53	5	433	58	491	11.81%	89.21%	8.62%	100.00%	54.31%	-4.13%	-16.29%
Average	1173.923	5.692308	64.61538	78.92308	1179.615	143.5385	1323.154	10.85%	93.82%	48.87%	99.46%	74.16%	0.49%	3.56%

Conclusion

This data partitioning method has shown improvement on drug and property recidivism offense types. A substantial increase for modeling all recidivism and increased accuracy for both drug and property recidivism modeling is also shown in these results. Our models for person offending, however, show that overall nothing is gained by partitioning the data.

We can see some remarkably strong prediction for specific layers. For example, Table 2.27, Layer 29, shows a layer for which the sensitivity/accuracy differential is 23.26 percent. Looking at accuracy alone, there is a small gain over all data (98.66 percent vs. 93.33 percent); the improvement in sensitivity, however, is much greater (87.72 percent vs. 41.64 percent).

It is for drug offending, however, that the benefits of co-clustering are greatest. Looking at the accuracy column we see a modest gain from 92.28 percent to an average of 96.29 percent, but the average sensitivity figure of 80.05 is a considerable improvement over 51.96 percent found for the sample as a whole.

CHAPTER 3
SPATIAL CLUSTERING OF JUVENILE RECIDIVISM

INTRODUCTION

Classification of persons or cases is central to studies of behavior and to standardizing decisions made about persons. The purposes of classification, at least with regard to juvenile delinquents, include: (1) improving our understanding of delinquent behavior (Jurkovic and Dodge, 1977; Wattenberg, 1979; Megargee et al., 1979; Warren and Hindelang, 1979; Frechette and LeBlanc, 1980; Moffitt and Caspi, 2001), (2) improved matching of offenders to interventions (Hunt and Hardt, 1965; Warren, 1971, 1976; Sechrest, 1987; Harris, 1988; Brannon et al., 1989; Mezzich, Coffman and Mezzich, 1991; Van Voorhis, 1991; Palmer, 1992, 1984), (3) offender population management (Baird, 1986; Glaser, 1987; Dembo et al., 1994) and (4) risk prediction (Warren, 1965; Gottfredson, D., 1987; Gottfredson, S., 1987; Brennan, 1987; MacKenzie et al. 1988; Andrews et al., 1990; Palmer, 1992; Bonta, 1966). All decisions involve prediction, and classification enables decision makers to search for and manage information relevant to the decision being made.

An undifferentiated examination of delinquency patterns or of individual delinquents is likely to mask relevant information about who is positively or negatively affected by what. For example, Palmer (1992) argues that differences among offenders and their individual circumstances will affect responses to specific intervention methods. Thus, what works with one type of offender may not work with another. This notion of responsivity has led clinical researchers to examine ways in which to classify and predict intervention outcomes, taking into account such critical differences as maturity, gender, cultural background, histories of abuse and neglect, and mental health disorders (Bonta, 1996). Program evaluators have been cautioned to examine program outcomes differentially, looking at differences in outcome within clusters of person of the same type (Palmer, 1984,).

Differentiating among persons using individual-level attributes, however, misses the critical influences of environmental forces. Recent work has underscored the need to incorporate ecological data into our classifications of individuals (Turner, Hartman and Bishop, 2007; Onifade et al., 2008). Doing so recognizes the mechanisms by which perceptions, attitudes and behavior patterns of individuals are shaped by experiences in specific environments. We also argue that different individuals in different environments may require different explanations of delinquency.

Co-Clustering with Plaid

This chapter builds on the use of co-clustering discussed in Chapter 2. Given the strong indication in our preliminary analyses of different causal models operating with different types of offenses, we decided to explore further the differential nature of recidivism by clustering cases using a novel approach to cluster analysis that involves the simultaneous clustering of subjects and subject attributes. The plaid model was fitted to the input data on juvenile recidivism. We did not include the response variable that specifically identifies juveniles who recidivate. The plaid model iterations terminated at 52 layers.

In the output, the number of rows and the number of columns identified for each layer are given. Also given for each layer is a value for mu (the overall layer effect), plus a value for each column (variable) effect included in that layer and a value for each row (juvenile) effect. Most layers have a large number of rows and a small number of columns. The number of juveniles in any layer ranges from 6 to 3357 and the number of variables in any layer ranges from 1 to 12. Some of the layers have small numbers of rows and so were deemed of little practical interest. We decided to restrict attention in the analysis only to those 13 layers that were identified by having at least 400 juveniles and at most 4 variables. We next give a brief description of the 13 layers. The recidivism rate over all 6,768 juveniles was 0.387.

The variables selected for the Plaid analysis are listed in Table 3.1.

Table 3.1

Variable Name	Description
sexoff	Was the initial crime a sexual offence?
WhiteDum	Is youth white?
HispanicDum	Is youth Hispanic?
Probation	Was youth on probation at the time of his arrest?
ParentalCrime	Does a parent have a criminal history?
LiveInstitution	Did the youth live in an institution at the time of the crime?
PriorPersonalChgs	Did the youth have prior personal offense charges at the time of the crime?
Juvdrgar	Did the youth have prior drug arrests at the time
	of the crime?
Prioroutofhomepl	Was the youth placed in out-of-home program at the time of the crime?

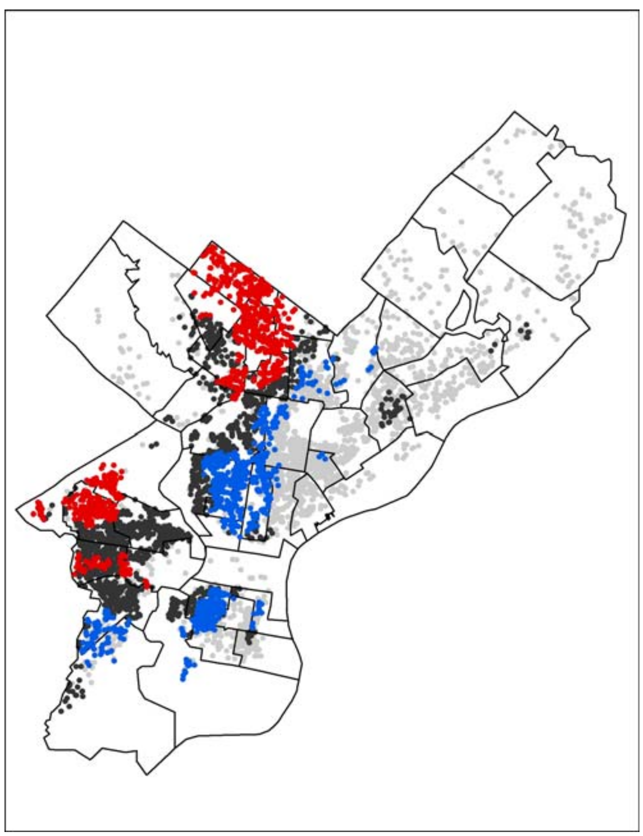
InstantPerson	Was the instant offense person-related?
InstantProperty	Was the instant offense property-related?
sibarr	Did any of the youth's siblings have an arrest record?
jhismh	Did the youth have a history of mental-health?
victinj	Did the youth injure a victim in the instant o&#11;ense?
age	How old was the youth at the case recording?
AlcoholAbuse	Did the youth have a history of alcohol-abuse?
DrugAbuse	Did the youth have a history of drug-abuse?
ParenDeceased	Is at least one parent deceased?
ParSubAbuse	Did the parents have a history of substance-abuse?
den dr sale	Density of drug sales near youth's home
den person	Density of person offenses near youth's home
p black	Percent black in census block
p vacant	Percent of vacant housing in census block
p spanish	Percent Hispanic in census block
p highsch	Percent high-school graduates in census block
kcnt 1km	Number of delinquent youths residing within 1 km of the juvenile
gi	Getis-Ord G. i statistic

For all derived layers, the Gi statistic was computed for each youth to see if there were spatial clusters of youths with strong or weak membership in the layer. This turned out to be very useful because when visualizing thousands of points, it is difficult to detect a spatial pattern visually. The maps of selected layers are given below. In each map, a youth point is colored red if there exists a significant local cluster of high degree of membership in the layer, blue if there is a significant local cluster of low degree of membership in the layer, dark gray if the youth is in the layer but not in a significantly high or low local cluster of membership, and light gray shows that the youth is not included in the layer.

Descriptions of Selected Layers

Layer 1 contains 3,357 juveniles and is identified by the 3 variables age, p_highsch, and p_black. Because mu = 1.31, these juveniles are about 1.31 standard deviations *above the mean* in these variables. The column (variable) effects range from –0.40 standard deviations for age to 0.60 standard deviations for percent black. The map for this layer shows clearly that the strongest membership for the layer is associated with Wynnefield and West Oak Lane, two primarily African-American neighborhoods that are generally working- and middle-class, as distinguished from many of the other African-American areas where the layer membership predominates and tend to be more socioeconomically disadvantaged. This pattern is confirmed by the mean values of the variables in this layer; in particular, the mean percent completing high school for this layer, 62.1%, is higher than the global mean of all youths in the study, 58.9%. The percent African American in this layer, 90.34%, is much higher than for the entire data set, 64.66%, and the average age of juveniles in this layer, 15.55, is slightly less than that for the entire data set. In other words, this layer captures a spatially clustered subgroup based upon race and educational attainment. Furthermore, the recidivism rate for the juveniles in this layer was 0.352, well below the rate over all juveniles.

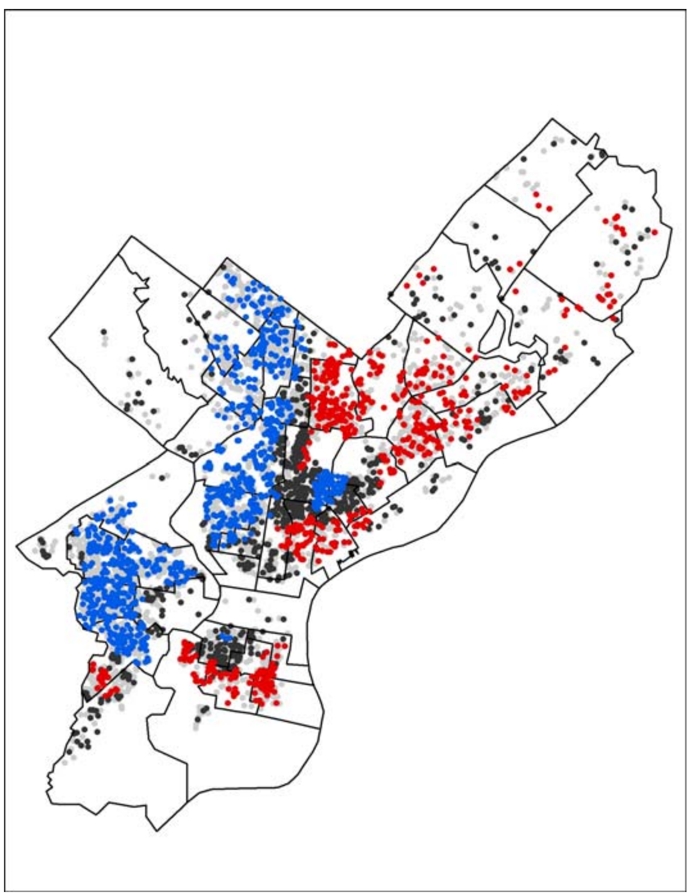
Layer 1



Layer 3 contains 2,279 juveniles and is identified by the 2 variables Age and DrugAbuse. Because mu = 1.21, the juveniles in this layer are 1.21 standard deviations *above the mean* on these variables. The column (variable) effect for Age is –0.64 standard deviations and the column effect for DrugAbuse is 0.64 standard deviations. All juveniles who are members of this layer had a history of drug abuse (as opposed to 46.34% for the entire data set) and the average age was 16.24 years, older than that for the entire data set.

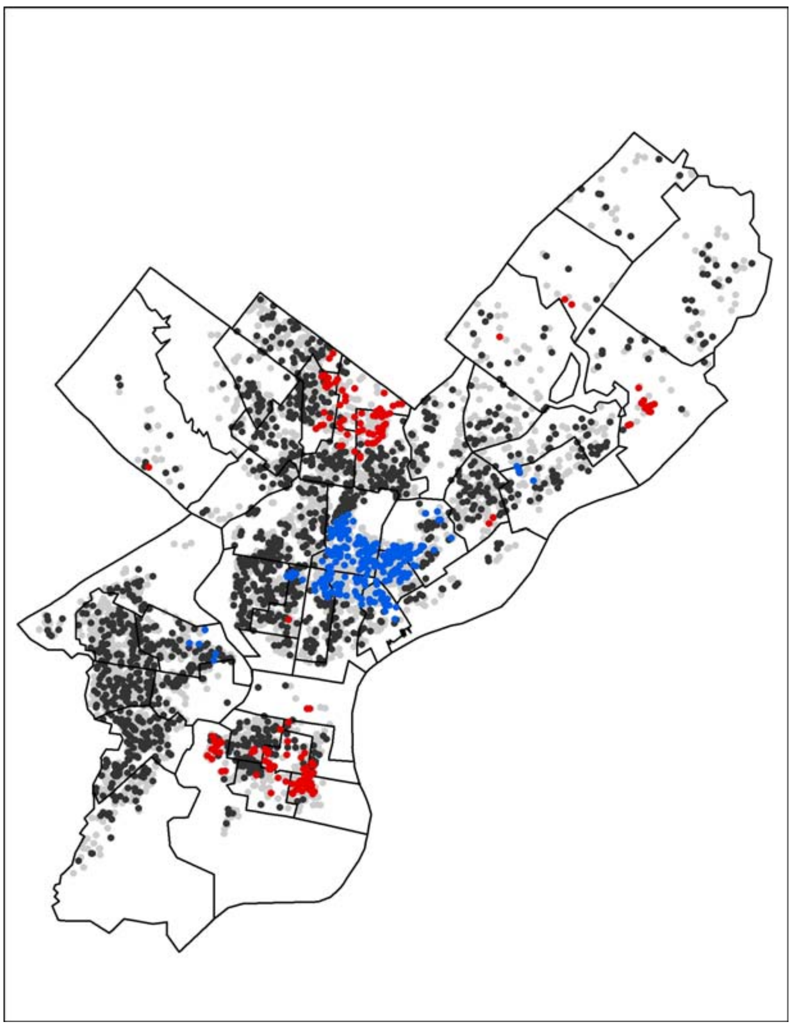
Unlike in Layer 1, the map of the cases in Layer 3 does not visually suggest strong spatial clustering. However, degree of membership is spatially clustered, with the strongest membership occurring in primarily white neighborhoods (such as the Northeast) and neighborhoods with substantial white populations, such as in southwest Philadelphia, parts of South Philadelphia, and in the Fishtown/Kensington/Port Richmond neighborhoods. Lower degree of membership is concentrated in areas dominated by African American and Hispanic residents with few whites. The recidivism rate for the juveniles in this layer was 0.427, much higher than the rate over all juveniles.

Layer 3



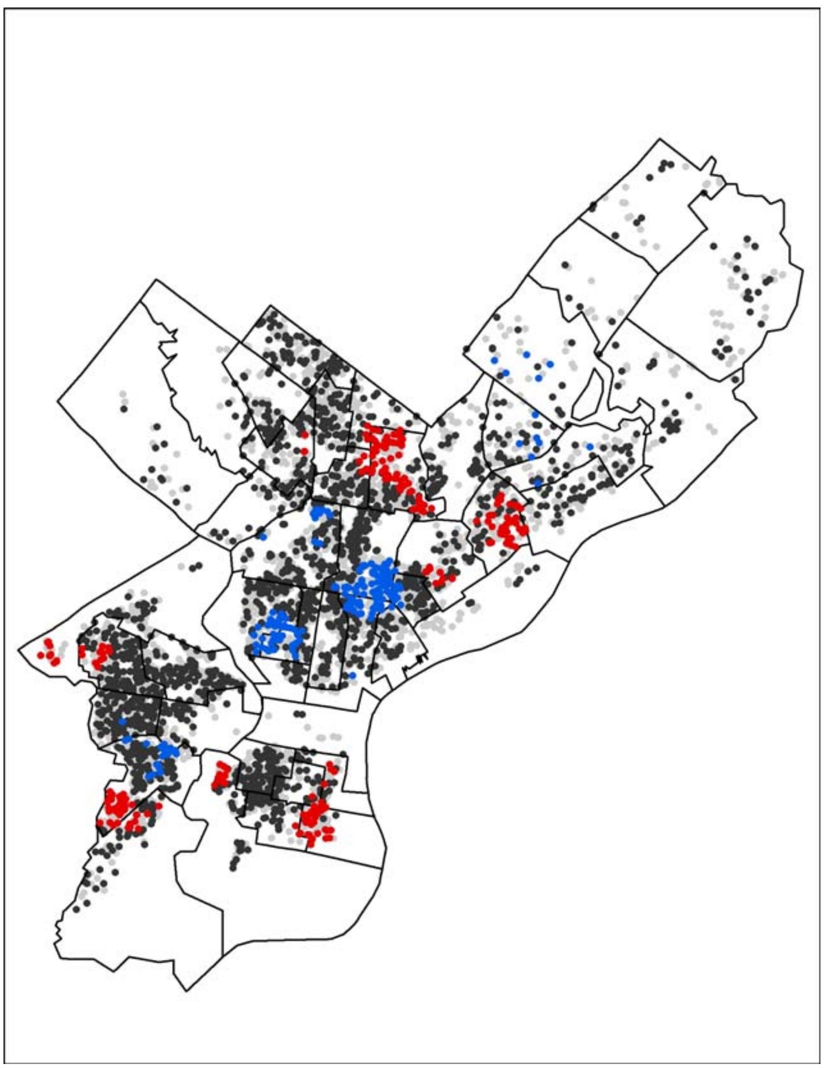
Layer 5 contains 1,989 juveniles and is identified by the 2 variables LiveInstitution and InstantProperty. Because mu = 1.14, the juveniles in this layer are 1.14 standard deviations *above the mean* on these variables. The column (variable) effect for LiveInstitution is −0.98 standard deviations and the column effect for InstantProperty is 0.98 standard deviations. All juveniles in this layer had a property offense as the instant offense and 33.68% lived in an institution, slightly higher than the rate for the entire data set. High degree of membership in Layer 5 occurs primarily in South Philadelphia and in the East Oak Lane and Olney neighborhoods. These neighborhoods are mostly working class and have a mix of white and African American residents. Low degree of membership occurs in the Hispanic Hunting Park neighborhood. The recidivism rate for the juveniles in this layer was 0.401, much higher than the rate over all juveniles.

Layer 5



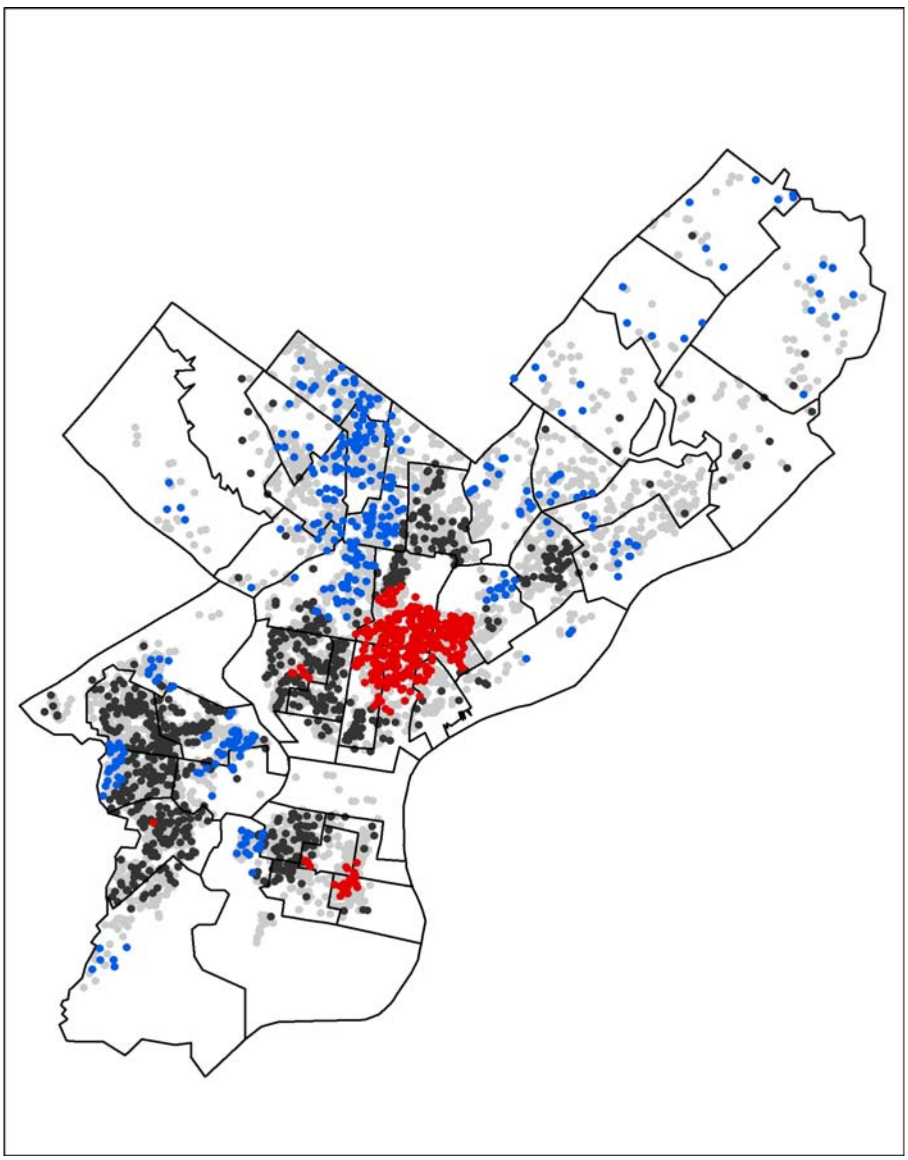
Layer 6 contains 2,385 juveniles and is identified by the 2 variables victinj and InstantPerson. Because mu = 1.28, the juveniles in this layer are 1.28 standard deviations *above the mean* on these variables. The column (variable) effect for victinj is –0.72 standard deviations and 0.72 for InstantProperty. All juveniles in this layer had a person offense as the instant offense and 45.53% of them injured a victim during the instant offense, much higher than the rate for the entire data set, 16.74%. Layer 6 exhibits strong clustering in high degree of membership in several disparate neighborhoods that have one specific characteristic in common – a mix of white and African American residents. This is relatively unusual in a very segregated city such as Philadelphia, but occurs in Wynnefield and southwest Philadelphia, Grays Ferry and Pennsport in South Philadelphia, and Olney and Frankford in northern Philadelphia. The recidivism rate for the juveniles in this layer was 0.334, close to the lowest rate among all the layers detailed here.

Layer 6



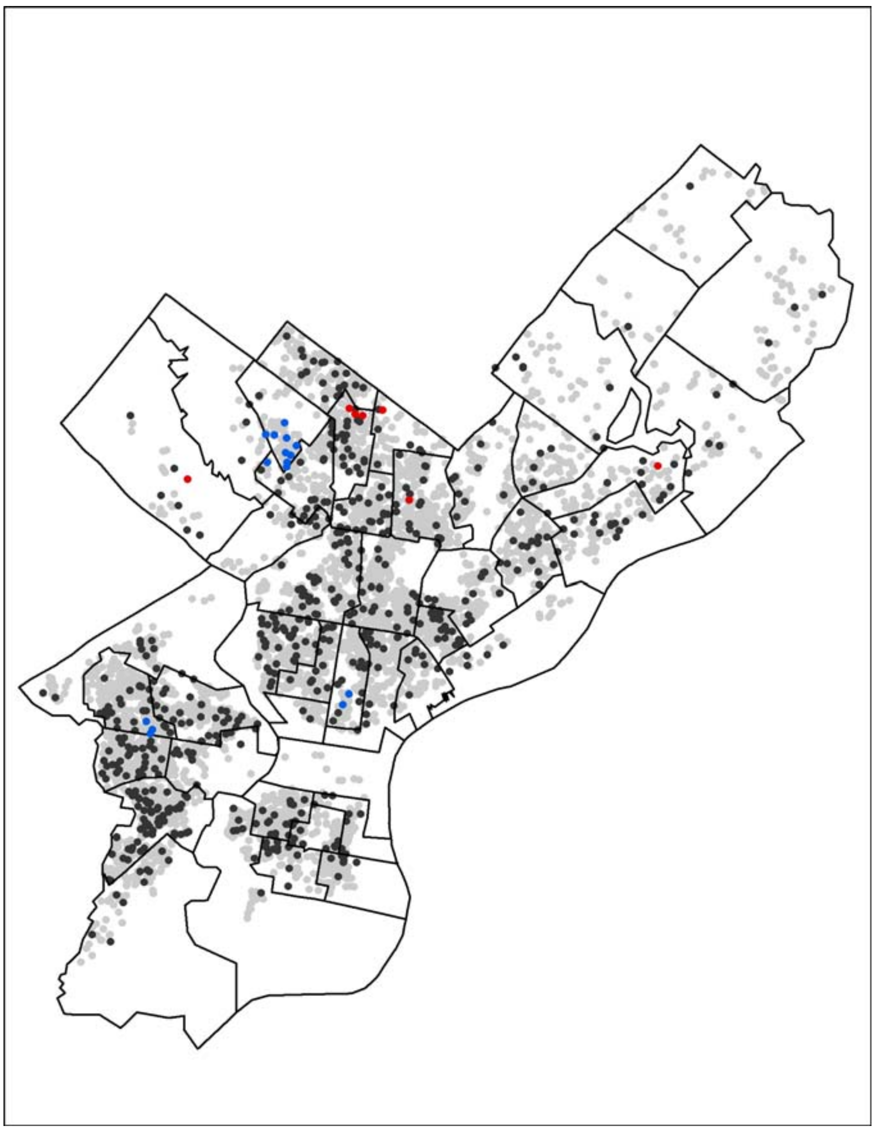
Layer 8 contains 1,657 juveniles and is identified by the 3 variables kcnt_1km (number of delinquent youths within 1 km), gi, and Juvdrgar. Because mu = 1.02, the juveniles in this layer are just over one standard deviation above the mean on these variables. The column effect for kcnt_1km is −0.53 standard deviations, for gi is −0.47 standard deviations, and for Juvdrgar is 0.99 standard deviations. All juveniles in this layer had a prior drug arrest, and the two "density" variables, kcnt_1km and gi, were both much higher than those values for the entire data set. Layer 8 exhibits high degree of membership in Hunting Park and the surrounding neighborhoods with large Hispanic populations. The recidivism rate for the juveniles in this layer was 0.460, the second highest such rate among all the layers detailed here.

Layer 8



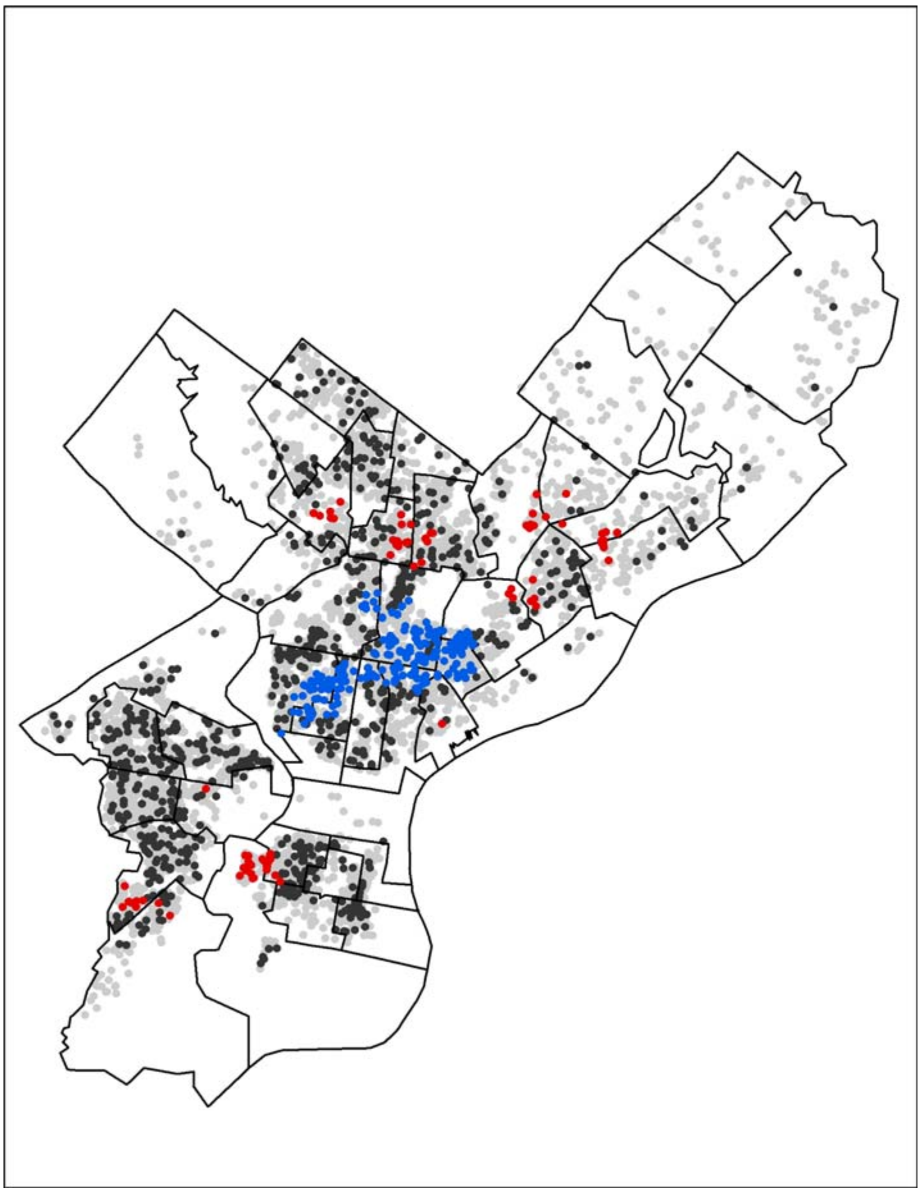
Layer 15 contains 641 juveniles and is identified by the 2 variables AlcoholAbuse and ParSubAbuse. Because mu = 1.04, the juveniles in this layer are just over one standard deviation above the mean on these variables. The column effect for AlcoholAbuse is –0.85 standard deviations and is 0.85 standard deviations for ParSubAbuse. All juveniles in this layer had a parent with a history of substance abuse and 26.52% had a history of alcohol abuse, higher that the rate for the entire data set, 21.14%. Layer 15 is not only in small in number of cases, but those cases are spread out throughout the city. There is little evidence of spatial clustering in degree of membership. The recidivism rate for the juveniles in this layer was 0.410, well above the rate over all juveniles.

Layer 15



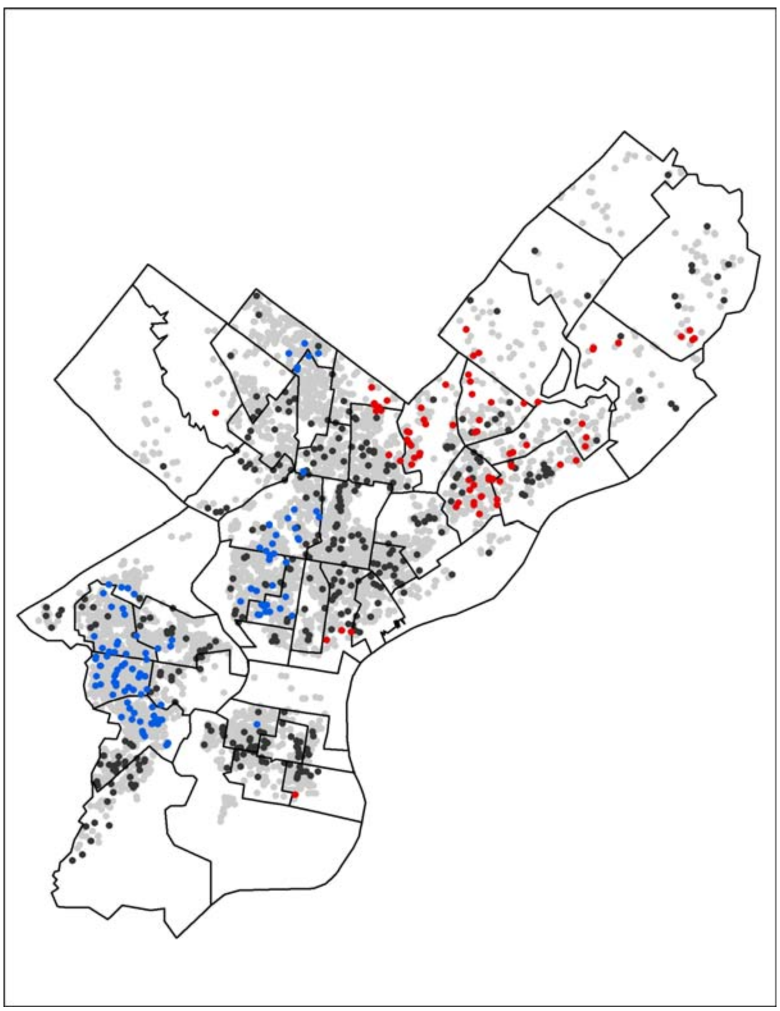
Layer 18 contains 1,172 juveniles and is identified by the single variable sibarr. Because mu = 1.95, the juveniles in this layer are 1.95 standard deviations above the mean on this variable. The column effect for sibarr is essentially zero: all juveniles in this layer had a sibling with an arrest record, compared with 25.74% for the entire data set. Layer 18 exhibits a handful of clusters of high degree of membership in neighborhoods with combinations of white and African American residents: Gray's Ferry, around Frankford, and Olney. The recidivism rate for the juveniles in this layer was 0.388, very close to the rate over all juveniles.

Layer 18



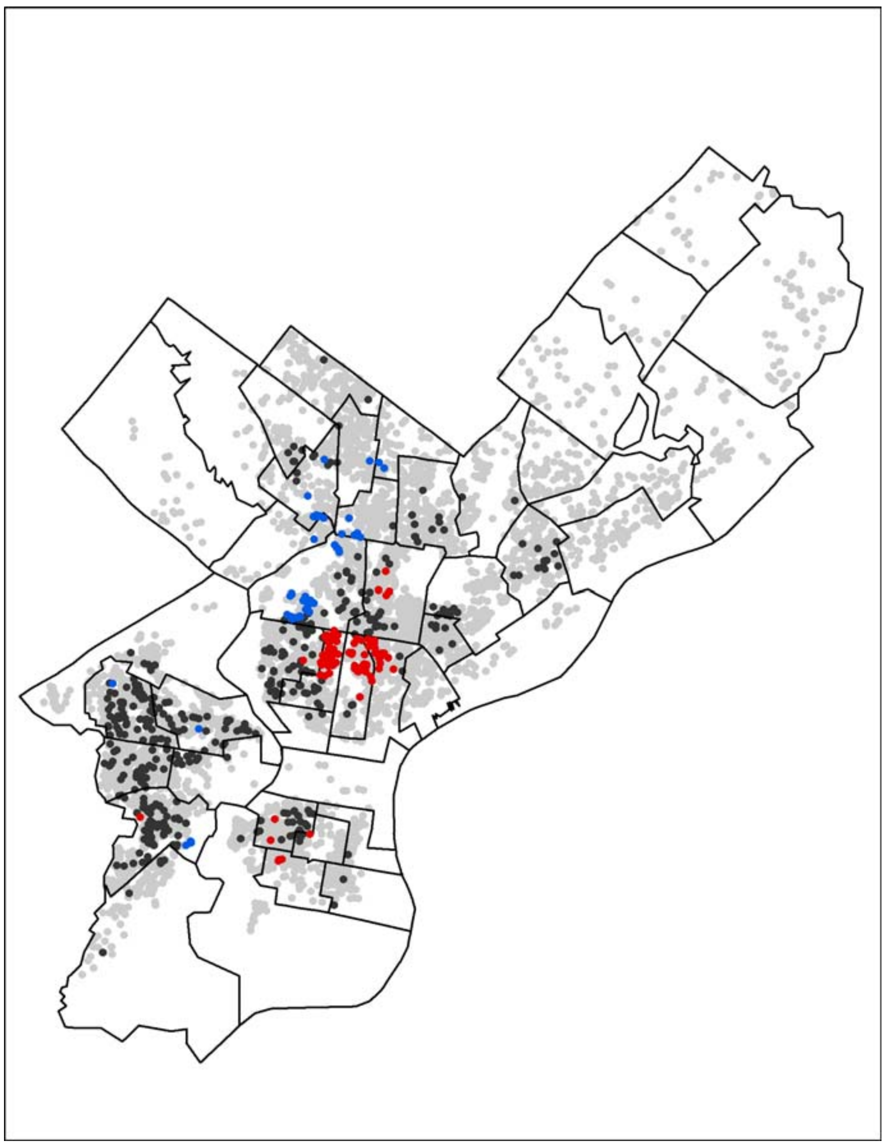
Layer 22 contains 563 juveniles and is identified by the 3 variables age, p_highsch, and PriorPersonalChgs. Because mu = 0.89, the juveniles in this layer are 0.89 standard deviations above the mean on these variables. The column effect for age is −0.57 standard deviations, for p_highsch is −0.36 standard deviations, and for PriorPersonalChgs is 0.93 standard deviations. Almost all juveniles (98%) in this layer had prior personal charges, their average age was 15.81 years (slightly higher than that for the entire data set), and 63.25% were high-school graduates (much higher than the 58.9% for the entire data set). Layer 22 indicates high degree of membership in the lower Northeast around the neighborhoods of Lawncrest, Oxford Circle, and Frankford. These are primarily white, working class neighborhoods, though Frankford also has a substantial African American population. Low degree of membership is observed in West Philadelphia and Strawberry Mansion/Nicetown neighborhoods of North Philadelphia, which are tend to be poor, with high concentrations of African American residents. The recidivism rate for the juveniles in this layer was 0.455, well above the rate over all juveniles.

Layer 22



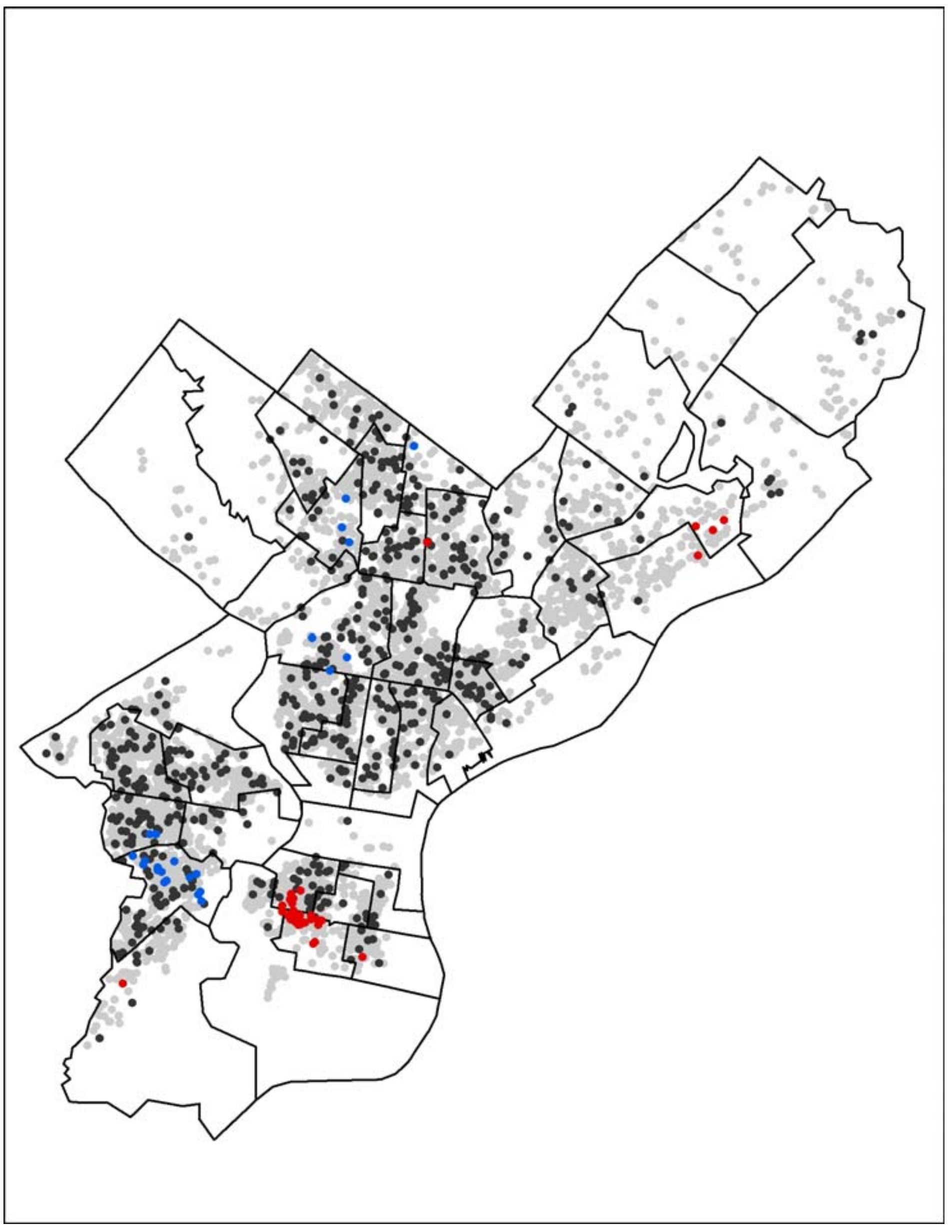
Layer 29 contains 536 juveniles and is identified by the 4 variables p_black, gi, kcnt_1km, and den_person. Because mu = 0.65, the juveniles in this layer are 0.65 standard deviations above the mean on these variables. The column effect for p_black is –0.20 standard deviations, for gi is 0.04 standard deviations, for kcnt_1km is 0.07 standard deviations, and for den_person is 0.08 standard deviations. For the juveniles selected for this layer, a high proportion, 85.66%, are African American, and the three "density" variables have much higher values than for the entire data set. Cases in Layer 29 occur almost exclusively in poor African American neighborhoods. The highest degree of membership occurs in North Philadelphia. The recidivism rate for the juveniles in this layer was 0.330, the lowest rate among all the layers detailed here.

Layer 29



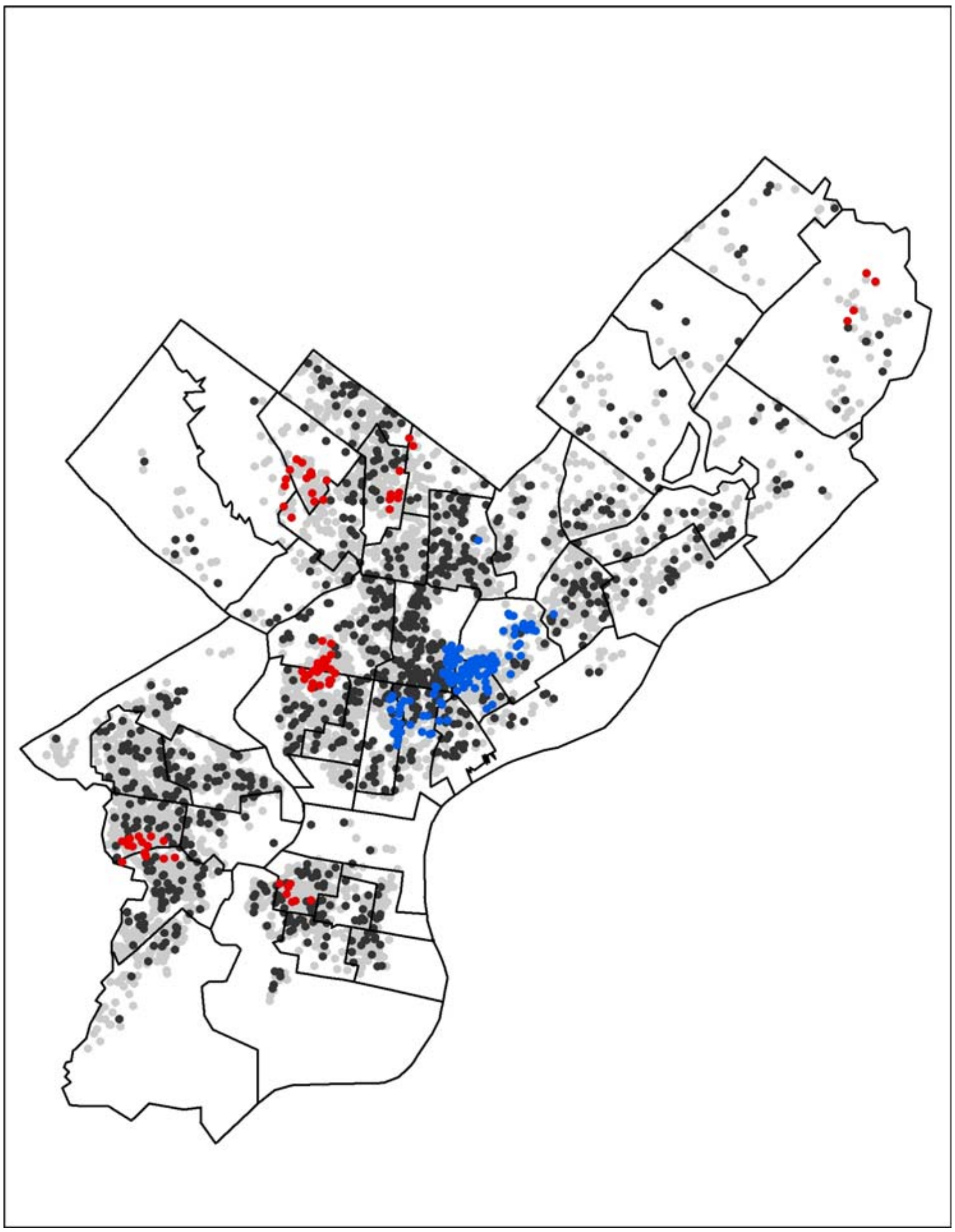
Layer 31 contains 697 juveniles and is identified by the single variable LiveInstitution. Because mu = 1.52, the juveniles in this layer are 1.52 standard deviations above the mean on these variables. The column effect for LiveInstiution is essentially zero. All juveniles in this layer lived in an institution at the time of their instant offense. Cases in Layer 31 occur primarily in African American and Hispanic neighborhoods of different classes. The highest degree of membership is clustered in South Philadelphia. The recidivism rate for the juveniles in this layer was 0.428, well above the rate over all juveniles.

Layer 31



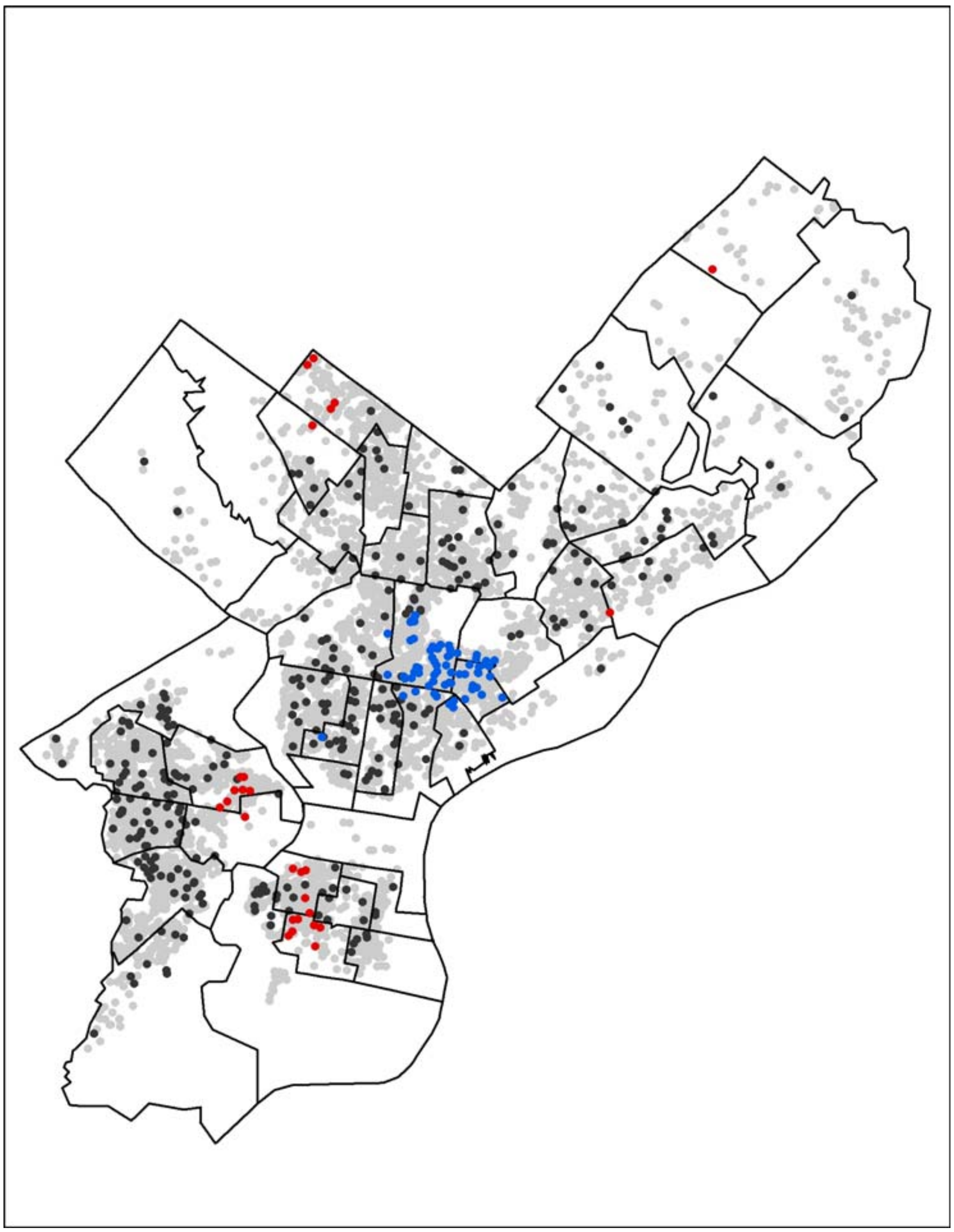
Layer 34 contains 1,202 juveniles and is identified by the 2 variables DrugAbuse and AlcoholAbuse. Because mu = 0.93, the juveniles in this layer are 0.93 standard deviations above the mean on these variables. The column effect for DrugAbuse is –0.68 standard deviations and is 0.68 standard deviations for AlcoholAbuse. All juveniles in this layer had a history of alcohol abuse and 87.19% of them had a history of drug abuse, compared with 46.34% for the entire data set. Several clusters of high degree of membership in Layer 34 occur in several different African American neighborhoods. The recidivism rate for the juveniles in this layer was 0.423, well above the rate over all juveniles.

Layer 34



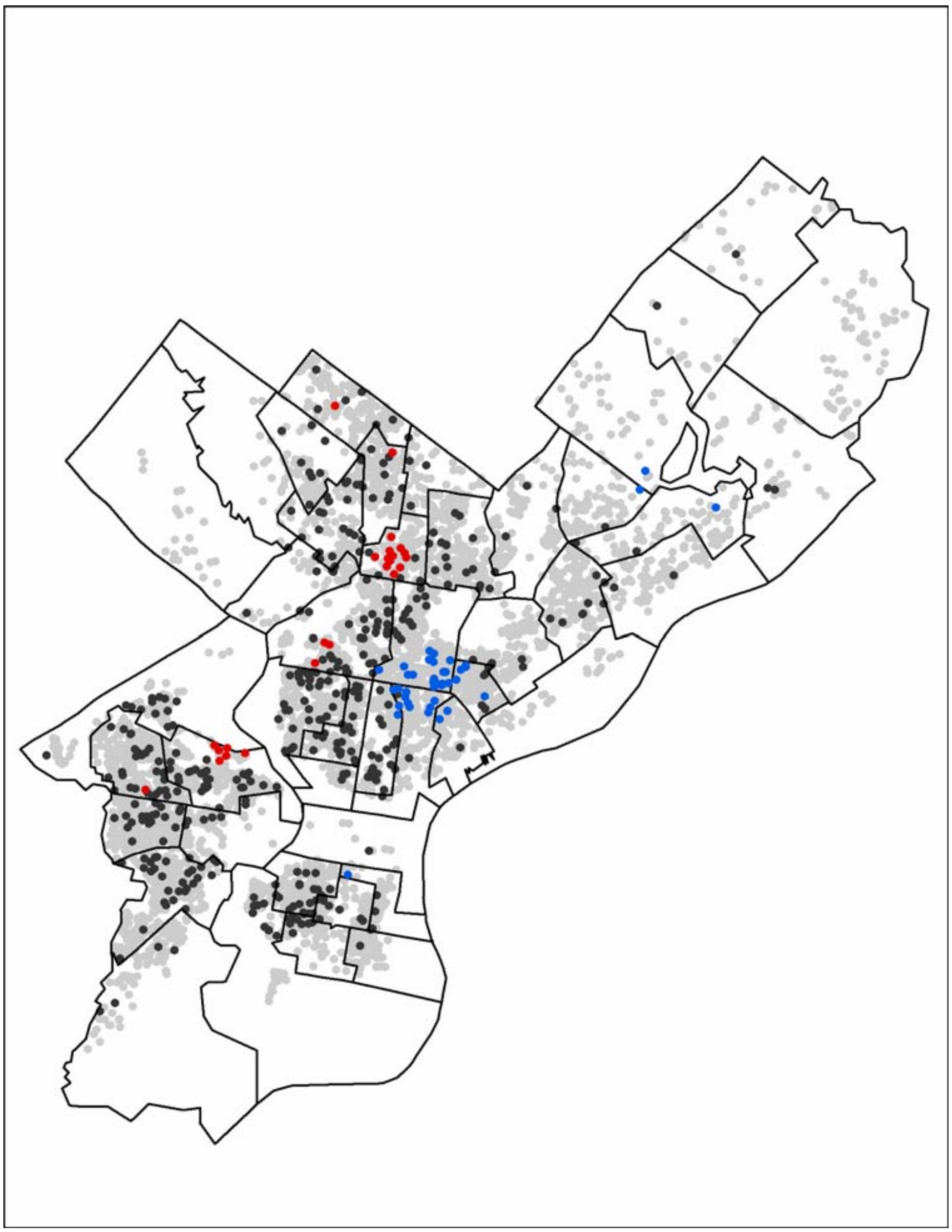
Layer 43 contains 424 juveniles and is identified by the single variable Probation. Because mu = 1.80, the juveniles in this layer are 1.80 standard deviations above the mean on these variables. All juveniles in this layer were on probation at the time of their instant offense, compared with only 9.47% in the entire data set. Cluster of high degree of membership for Layer 43 occur in Point Breeze, Mantua, and West Oak Lane. These are all African American neighborhoods. The recidivism rate for the juveniles in this layer was 0.481, the highest rate of all the layers detailed here.

Layer 43



Layer 50 contains 496 juveniles and is identified by the single variable PriorPersonalChgs. Because mu = 1.75, the juveniles in this layer are 1.75 standard deviations above the mean on this variable. The column effect for this variable is essentially zero. All juveniles in this layer had prior personal charges, compared with only 19.15% in the entire data set. Layer 50 contains cases primarily from African American neighborhoods. The recidivism rate for the juveniles in this layer was 0.450, well above the rate over all juveniles.

Layer 50



Discussion

What is particularly notable about the results of the Plaid analysis is the spatial dependency of the layers and the related differences in recidivism rates. Table 3.2 below summarizes the defining variables, spatial concentrations and recidivism rates for each layer.

Table 3.2.

Layer	Defining Variables	High Degree Neighborhoods	Low Degree Neighborhoods	Recidivism Rate
1	p_highsch-high p_black-high	Wynnefield West Oak Lane	Poplar, Northern Liberties	Low (.35)
3	Age – high DrugAbuse - high	Southwest Philadelphia South Philadelphia Fishtown, Kensington, Port Richmond	Hunting Park West Phila/Strawberry Mansion	High (.43)
5	LiveInstitution - high Instant Property - high	South Philadelphia Oak Lane, Olney	Hunting Park	High (.40)
6	Victinj Instant Person	Wynnefield Southwest Phila Greys Ferry/Pennsport Olney/Frankford	Starwood Hunting Park	Low (.33)
8	Kcnt_1km Gi juvdrugar	Hunting Park	Logan Ogontz	V. High (.46)
15	AlcoholAbuse ParSubAbuse	No pattern	No pattern	High (.41)
18	Sibarr	Gray’s Ferry Frankford Olney	Hunting Park Strawberry Mansion	Avg. (.39)
22	P_highsch (high) PriorPersonalChgs (high)	Lower Northeast: Lawncrest, Oxford Circle, Frankford	West Phila: Strawberry Mansion, Nicetown	V. High (.46)
29	P_black (high) Gi (high) Kcnt_1km (high) Den_person (high)	North Philadelphia: Starwood, Poplar	Nicetown/Tioga	V. low (.33)
31	LiveInstition (all)	Girard Estates	Kingsessing	High (.43)
34	DrugAbuse (high) AlcoholAbuse (high)	Several African American neighborhoods	Hunting Park	High (.42)
43	Probation (all)	Point Breeze Mantua West Oak Lane	Hunting Park	V. High (.48)
50	PriorPersonalChgs (all)	Logan Mantua	Hunting Park	High (.45)

Those layers with high or very high recidivism rates are associated with high concentrations of youths who where on probation at the time of their offense, had prior personal charges, were on aftercare, used drugs and were older. In only one layer, Layer 15, did parental substance abuse appear, and it appeared with alcohol abuse on the part of the youth. Members of this layer were not spatially concentrated but instead were spatially ubiquitous.

Neither p_highsch or p_black alone were associated with high recidivism rates, but in combination, they were associated with low recidivism. Thus, while we found several predominantly African American neighborhoods with concentrations of recidivists, when the neighborhood also had a higher than average number of persons with high school educations, recidivism rates were lower than average.

Hunting Park, a neighborhood that is largely Hispanic, was notable in its association with low membership in 7 of the 11 layers. This neighborhood was associated only with Layer 8, a layer defined by a high count of other delinquent youths and a high number of youths with prior drug arrests, and a layer with a high rate of recidivism.

Individual-level attributes also play an important role in this analysis. Layer 3 consists of youths who are older and who are known to abuse drugs. These youths reside in areas that are are economically disadvantaged and comprised of largely African American populations. The combination of age and drug abuse is associated with a high rate of recidivism.

Another layer with a high recidivism rate consisted of property offenders who are on aftercare, meaning that they had just reentered the community following a period of incarceration, were split between South Philadelphia and two neighborhoods in the Northwest section of the city. Populations in these neighborhoods are largely working-class African Americans.

In both layers with Prior Person Charges (Layers 22 and 50) had high rates of recidivism (.46 and .45, respectively), but were most highly clustered in different neighborhoods. Layer 50 is most highly concentrated in the Logan area of North Philadelphia, an African American neighborhood, but least highly concentrated in Hunting Park, a largely Hispanic area. Layer 22, however, is concentrated in Northeast Philadelphia, a racially mixed part of the city, and least concentrated in West Philadelphia, another racially mixed area.

Interestingly, the layer most associated with violent offending (a violent instant offense and injury to the victim) was associated with a low recidivism rate. This layer 's neighborhoods are best characterized as racially (mainly White and African American) and economically mixed (poor and working class).

What we see in this analysis is that different combinations of persons and their individual and environmental attributes are associated with different neighborhoods, and that these clusters of persons and attributes suggest different causes of recidivism. Small differences in educational level and economic advantage are associated with different types of offending and different levels of reoffending. In addition, risk factors such as alcohol and drug abuse, including parental drug abuse, and prior involvement with the justice system increase the likelihood that layer properties are associated with high recidivism.

CHAPTER 4
HIERARCHICAL LINEAR MODELING OF JUVENILE RECIDIVISM

INTRODUCTION

Hierarchical linear models (hlm) rose to prominence over the last decade as a methodology of choice for naturally nested data. The youth in the PRODES dataset are nested simultaneously in the neighborhoods they live and programs they were assigned to by the court. This dual nature of nesting lends itself to a newer technique in hlm called the cross-classified model. Figure 4.1 provides a graphical representation of the nesting structure of the data to be analyzed.

Figure 4.1: Cross-Classified Nesting Structure of Data

	Program 1	Program 2	Program 3	Program 4	Program 5
Neighborhood 1	12	15	9	32	12
Neighborhood 2	43	8	31	27	17
Neighborhood 3	19	36	28	14	4
Neighborhood 4	16	21	26	53	33
Neighborhood 5	31	11	0	18	22

Figure 4.1 illustrates a matrix in which the hypothetical values in the cells represent the count of juveniles that are nested within each neighborhood and program. For example, there are 12 juvenile offenders who live within Neighborhood 1 and attended Program 1.

We ran descriptive statistics, exploratory crosstabs and correlations to investigate the relationship of the independent variables to the outcome variables. Following the descriptive analyses, we ran preliminary traditional logistic regression analyses to further understand the relationships between the independent variables and the four outcomes. We began the hlm models by examining two, traditional, two-level models 1) youth nested in neighborhoods and 2) youth nested in programs. The two-level models were followed by a cross-classified model which allows for estimation of neighborhood effects on recidivism controlling for program effects and program effects on recidivism, controlling for neighborhood effects. This cross-classified model, like the individual two-level models simultaneously controls for the individual level characteristics.

A cross-classified generalized linear model allows for estimation of very complex research questions. As mentioned in previous chapters, we were exploring the simultaneous effects of neighborhoods and programs while controlling for the individual youth history.

The simplest form of a hierarchical model is a two-level model where subjects are nested within one level. A more complex model would be a three-level model, where subjects are nested within level-2 and where level-2 units are nested within level 3. At first it may appear that a three-level model is appropriate for youth nested within programs within neighborhoods. This is not the case however, because there would have to be large numbers of youth within each program and each program would need to be represented within multiple neighborhoods. In the city of Philadelphia, each program appears in only one neighborhood, thus the cross-classified model suits the data structure well. Sparse cell counts are of no concern because marginal totals are used in the estimation process, whereby the neighborhood effects are estimated across programs and programs effects are estimated across neighborhoods.

The level-1 unit of analysis for this study is youth, and the binary outcome for level-1 is whether or the youth recidivates with a particular crime (0=no, 1=yes) within a six month follow-up period after their time in treatment. Similar to traditional logistic regression analyses, a logit link function will be used to model the binary outcome at level-1. Therefore the outcome will be $\eta_{ijk} = \log(\varphi_{ijk}/1-\varphi_{ijk})$, where η_{ijk} is the log odds recidivism. Using the regression coefficients, we can take $e^{\eta_{ijk}}$ to convert the predicted log odds to a predicted odds. We can also convert the predicted odds to a predicted probability for any case using

$\varphi_{ikj} = 1/\ (1 + e^{-\eta_{ijk}})$.

At level-1, the model looks like:

$$\log(\varphi_{ijk}/(1-\varphi_{ijk})) = \pi_{0jk} + e_{ijk}, \qquad (4.1)$$

$e_{ijk} \sim \varphi_{ijk} * (1-\varphi_{ijk})$

where:

$\log(\varphi_{ijk}/1-\varphi_{ijk})$ = log odds of recidivism of youth i in neighborhood j in program k;

π_{0jk} = the mean log odds of youth in cell jk, youth in neighborhood j and program k;

e_{ijk} = the random “youth effects”, the deviation of a youth’s ijk’s log odds from the cell mean.

The indices i,j,k refer to youth, neighborhoods and programs respectively.

The level-2 units of analyses here are the neighborhoods cross-classified by programs. Variation between cells can be attributed to neighborhood effects, program effects or a neighborhood by program effect. Here π_{0jk}, the mean log odds for youth in cell j,k, is modeled at level 2:

$$\pi_{0jk} = \theta_0 + b_{00j} + c_{00k} + d_{0jk} \qquad (4.2)$$

$b_{00j} \sim N(0, \tau_{b00})$,

$c_{00k} \sim N(0, \tau_{c00})$,

$d_{0jk} \sim N(0, {}_{td00})$,

where:

θ_0 is the grand mean log odds of recidivism for all youth;

b_{00j} is the random main effect of neighborhood j, the contribution of neighborhood j averaged over all programs;

c_{00k} is the random main effect of program k, the contribution of program k, average over all neighborhoods;

d_{0jk} is the random interaction effect, the deviation of the cell mean log odds from that predicted by the grand mean log odds and the two main effects log odds (a specific neighborhood and program).

If we substitute equation 4.2 into 4.1 yields a combined model

$$\log(\varphi_{ijk} / (1-\varphi_{ijk})) = \theta_0 + b_{00j} + c_{00k} + d_{0jk} + e_{ijk} \qquad (4.3)$$

The first step in an HGLM analysis is to assess whether or not the level-1 outcome varies across level-2 units. In this case, does youth recidivism vary across neighborhoods and/or programs. In hierarchical linear models in general, this is done by estimating an unconditional model. An unconditional model is a model that estimates the mean outcome with no predictors at any level. Estimating the unconditional model allows one to partition the variance in the outcome into distinct components. In this study, for example, one can estimate the variation that lies between neighborhoods, between programs, and—given a large enough sample size in each cell – a program within a specific neighborhood. For the two-level, cross-classified model, the total variance in the outcome is partitioned into within- cell components ($\varphi_{ijk} * (1-\varphi_{ijk})$) and between-cell components. Between-cell components are partitioned further into variance between neighborhoods (b_{00j}), variance between programs (c_{00k}) and residual variance, (d_{0jk}), unexplained by either neighborhood or program effects.

In traditional two-level hierarchical models with linear outcomes, the intraclass correlation coefficient is estimated in the unconditional model to identify whether or not the level-1 outcome varies across level-2 units. In the case of the cross-classified model, there are two interesting intraunit correlation coefficients: 1) the correlation between outcomes of two youth who live in the same neighborhood but are assigned to two different programs, 2) the correlation between the outcomes of two youth who are assigned to the same program but live in two different neighborhoods.

The unconditional model is less helpful in assessing whether the level-1 outcome varies across groups in cases with a binary outcome, because the variance is heteroscedastic. A crude approach would be to assume the Bernoulli outcome is normally distributed and compute the intraclass correlation coefficient as usual. This approach approximates the partitioning in the variance; however, it is unreliable for extreme probabilities (Goldstein, Browne & Rasbash, 2003).

We chose a fairly simple, intuitive method for this study. The method suggested by Raudenbush & Bryk (2002) is to visually inspect the interval of plausible values for b_{00j} (the neighborhood effect on the mean log odds for the outcome) and c_{00j} (the program effect on the mean log odds for the outcome). These intervals are computed as such: estimated b_{00j} from the unconditional model plus or minus the estimated τ_{b00} or τ_{c00} respectively. A small interval of plausible values for b_{00j} indicates that the rate of repeated mean log odds across neighborhoods does not vary much. One would then choose to fix (as oppose to allow to vary) the parameters for neighborhoodal effects. Likewise, a small interval of plausible values for c_{00j} indicates that the rate of repeated mean log odds across programs does not vary much. One would then choose to fix the parameters for program effects. In the case of no neighborhood or program effects, there would be no need to use an HGLM. One could revert to a traditional fixed coefficient logistic regression model.

The unconditional model also gives us the average log odds of youth recidivism at a typical neighborhood (θ_0). Transforming that value, we get the average probability of youth recidivism in a typical neighborhood, which in our case was .31. The interval of plausible values for the probability of youth recidivism across neighborhoods was (.390, .413), indicating that there is variance across neighborhoods. Similarly, for programs the interval of plausible values for the probability of youth recidivism was (.358, .381). Note the variance across neighborhoods is larger than the variance across programs. This is an indication that program characteristics are more predictive of youth recidivism than are neighborhood effects. The corresponding intervals for recidivism across neighborhoods by offense type differ by drug (.126, .142), person (.094, .108), and property recidivism (.104, .119). Similarly, the intervals for recidivism across programs differ by drug (.108, .124), person (.094, .108) and property recidivism (.097, .112).

Deeming the cross-classified HGLM appropriate for the analysis, we proceeded to build the models. To build the conditional model, we started with the level-1 only model and entered all eligible level-1 independent variables into the model. Variables with relationships to the outcome that varied across neighborhoods and/or programs, which we thought we would model further on in the analysis, were given random variance components. All other variables, including dummy variables, were grand mean-centered and given fixed variance components, making the interpretation of the grand mean the average log-odds of recidivism for a “typical” youth.

Following the level-1 only model, we built a level-2 only model. This level-2 only model allowed us to understand what the neighborhood and program effects were uncontrolled for youth traits. Following the level-2 only model, we built an *intercept as an outcome model* using only neighborhood variables, followed by an *intercept as an outcome model* that included only program variables.

Finally, we built a combined cross-classified model to simultaneously control for neighborhood and program contexts.

Findings of HLM Studies

<u>**Two-Level Analysis: Individual and Neighborhood**</u>

Dependent Variables

Our study examined four recidivism outcomes: any charge for a new offense, an aggregate measure, and three specific types of offense (property, drug, or person). Our decision to examine offense types separately was driven by preliminary analyses showing spatial patterns of offense types that suggested different causal mechanisms. Cases were followed into the adult criminal justice system. We conceptualized recidivism by using a dichotomous variable that measured whether or not a juvenile had a new petition filed against them at any time during their stay in the community-based program through a six-month post-program time period (0= no new petition filed, 1= new petition filed). Using new petitions to court as the measure of recidivism is most similar to the conceptualization of adult recidivism as a referral to court. Snyder & Sickmund (2006) found that using referral to court to measure recidivism produced an aggregate recidivism rate from several states of forty-five percent; a rate that they found lies between recidivism rates using rearrest and those using reconviction and reincarceration/reconfinement. Of our subjects, slightly more than forty percent had a new petition filed against them during the study period, a finding similar to the statewide statistics that use referral to court as their indicator for recidivism.

The period of study for which the juvenile is eligible to recidivate includes the time that they are in the community-based program and six-months following their release from their program. The average length of time per juvenile spent in the community-based program was 203 days, or slightly under seven months. Therefore, this study includes an average period of more than thirteen months from the point of disposition to six months after the date of release from a community-based program, during which time these youths juveniles were exposed to home and neighborhood influences. Outcome measures representing re-offending in- and post-program were available, but were not included as models in this study due to a similarity in significant effects and overall rates. Twenty-four percent of the juveniles in the dataset reoffended while attending a program, while twenty-two percent reoffended post-program.

Individual-Level Predictors

Several preliminary analyses, including binary logistic regression, CHAID (Chi-squared Automatic Interaction Detector), and neural network analyses, were conducted to pare down the total number of predictors in the ProDES database from several hundred to the thirty that appear in Table 4.1. Variables eligible for entry the current analysis were identified in the literature or were identified by one of the analyses mentioned above. They represent juvenile demographics, family traits, current offense characteristics, and criminal history. All of the individual-level variables selected are dichotomous (0=no, 1=yes) with the exception of the continuous variables of age and age at first arrest.

Neighborhood-Level Predictors

Data from the 2000 Decennial Census and the Philadelphia Police Department were used to construct indices representing neighborhood processes of disadvantage and crime. Five items from the Census 2000 Summary File 3 were extracted at the Census tract level, including: persons unemployed, female-headed households with children, persons living below the poverty line, persons on public assistance, and black residents. The raw counts per Census tract were aggregated to the 210 neighborhoods and then converted into proportions. We created an index of neighborhood disadvantage from a linear combination of the five items that exhibits a high level of internal reliability (alpha = .880). This operationalization of neighborhood disadvantage is consistent with studies of the effects of neighborhood context on crime (Baumer, Horney, Felson, & Lauritsen, 2003; Baumer, 2002; Morenoff, 2001; Sampson, Raudenbush, & Earls, 1997). Indicators representing racial heterogeneity and residential mobility, neighborhood processes identified by social disorganization theory to influence rates of crime, were constructed but excluded from the final models due to their lack of significant effect on the outcome measures.

The Philadelphia police data, represented by point-level locations of homicide, robbery, assault, theft, vehicle theft, burglary, and drug offenses in Philadelphia for a three-year period from 2000-2002, were aggregated at the 210 neighborhood level to create counts of crime per neighborhood. With the count of crime per neighborhood, the values were divided by the total population in the neighborhood to create a rate of each crime type, and then multiplied by 1000 to get a more meaningful value representing each crime type per 1000 residents in each of the neighborhoods.

Table 4.1 includes the dependent and independent variables considered for use in this analysis, along with the corresponding descriptive statistics.

Table 4.1: Descriptive Statistics of Individual- and Neighborhood-Level Variables

Outcome:	Metric	N	Mean	S.D.
Recidivism	0=no, 1=yes	7282	0.404	0.491
Recidivism: Drug Crime	0=no, 1=yes	7282	0.142	0.349
Recidivism: Person Crime	0=no, 1=yes	7282	0.101	0.301
Recidivism: Property Crime	0=no, 1=yes	7282	0.111	0.315
Level 1: Individual	Metric	N	Mean	S.D.
Sibling(s) Arrested	0=no, 1=yes	7282	0.263	0.440
Family Receives Public Assistance	0=no, 1=yes	7282	0.316	0.465
Parental Alcohol Abuse	0=no, 1=yes	7282	0.140	0.346
Parental Drug Abuse	0=no, 1=yes	7282	0.204	0.403
Parental Criminal History	0=no, 1=yes	7282	0.160	0.367
Parent(s) Deceased	0=no, 1=yes	7282	0.122	0.327
DHS Referral	0=no, 1=yes	7282	0.237	0.452
History of Family Violence	0=no, 1=yes	7282	0.105	0.306
Juvenile Age (years)	Continuous	7282	15.709	1.674
Age at 1st Arrest (years)	Continuous	7282	14.186	1.690
Juvenile Has Children	0=no, 1=yes	7282	0.040	0.192
White	0=no, 1=yes	7282	0.112	0.315
Black	0=no, 1=yes	7282	0.735	0.441
Hispanic	0=no, 1=yes	7282	0.131	0.338
Lives with Parent(s)	0=no, 1=yes	7282	0.524	0.499
Lives with Other Relatives	0=no, 1=yes	7282	0.087	0.281
Aftercare Case	0=no, 1=yes	7282	0.355	0.479
Prior Drug Offense	0=no, 1=yes	7282	0.330	0.470
Prior Personal Offense	0=no, 1=yes	7282	0.203	0.402
On Probation at Time of Arrest	0=no, 1=yes	7282	0.010	0.296
Prior Out-of-Home Placement	0=no, 1=yes	7282	0.076	0.265
Juvenile Alchohol Use	0=no, 1=yes	7282	0.221	0.415
Juvenile Drug Use	0=no, 1=yes	7282	0.478	0.501
Sex Offense	0=no, 1=yes	7282	0.059	0.236
Personal Offense	0=no, 1=yes	7282	0.317	0.465
Drug Offense	0=no, 1=yes	7282	0.235	0.424
Property Offense	0=no, 1=yes	7282	0.317	0.465
Victim Injured	0=no, 1=yes	7282	0.163	0.370
Weapon Involved	0=no, 1=yes	7282	0.145	0.352
Total # of Charges	Continuous	7282	4.876	3.987
Others Arrested for Offense	0=no, 1=yes	7282	0.401	0.491
Level 2: Neighborhood	Metric	N	Mean	S.D.
Disadvantage:*	Scale Item	210	0.000	0.822
Unemployment	Proportion	210	0.072	0.027

Poverty	Proportion	210	0.284	0.137
Public Assistance	Proportion	210	0.130	0.084
Female-Headed Households	Proportion	210	0.459	0.160
Black	Proportion	210	0.606	0.362
Crime	Rate per 1000 residents	210	219.948	111.213

* Cronbach's α = 0.880

Analysis

In addition to the univariate descriptive statistics, bivariate correlations of the individual-level variables were analyzed for associations and multicollinearity. Preliminary individual-level-only models identified fourteen variables to be used in further analyses. The selected variables are shown in Table 4.2 with frequencies by type of recidivism.

The frequencies of many of these variables are higher for juveniles who recidivated when compared to juveniles who did not recidivate, as would be expected of variables that have been identified by prior research as risk factors of juvenile recidivism. A comparison of these variables by recidivism type clearly indicates that juveniles in this population who recidivated via drug crimes are different from juveniles who recidivated by committing person or property offenses. Drug crime recidivists were much more likely to be Hispanic, have a prior drug arrest, and have had a prior out-of-home placement. An examination of the correlations and further follow up of the tolerance statistics indicated that there do not appear to be any issues of multicollinearity among the variables. Therefore, all of the indicators were eligible for inclusion in the analysis.

We followed the level-1 exploratory models with a two-level, hierarchical, generalized linear model to estimate individual and neighborhood effects on the odds of recidivism. We use a random intercept model to predict all four outcomes. The disadvantage and crime predictors at the neighborhood- level were entered in separate models for each outcome due to multicollinearity between the measures.

Table 4.2: Frequencies of Individual-level Variables by Recidivism Type

Predictor	Recidivism (n=4344)	Recid = Yes (n=2938)	Drug (n=1033)	Person (n=735)	Property (n=811)
White	12.05%	9.87%	6.10%	9.93%	13.70%
Hispanic	12.14%	14.61%	21.49%	9.39%	12.10%
Age at 1st Arrest	14.23	14.12	14.43	13.83	13.99
Parental Drug Abuse	18.89%	22.57%	21.39%	24.22%	22.32%
Parental Criminality	14.58%	18.11%	17.62%	20.44%	17.88%
Aftercare Case	30.10%	43.54%	48.49%	38.99%	40.50%
Public Assistance	29.40%	34.75%	36.88%	32.52%	34.65%
Sex Offense	7.68%	3.34%	2.13%	4.77%	3.34%
Personal Offense	39.04%	31.35%	23.62%	40.95%	32.43%
Property Offense	30.41%	33.70%	26.52%	32.65%	43.65%
On Probation	8.21%	11.83%	11.43%	11.87%	11.36%
Prior Drug Offense	28.94%	38.93%	59.59%	26.67%	24.66%
Prior Person Offense	17.50%	24.34%	22.33%	28.20%	22.08%
Prior Placement	5.26%	11.01%	32.07%	10.48%	10.77%

Results

Results of the eight hierarchical generalized linear models are shown in Table 4.3. Odds ratios, with their subsequent significance, are displayed. Odds ratios indicate the relative likelihood that a juvenile will recidivate, while holding all other predictors in the model constant. Most of the individual-level predictors in the models are dichotomous (indicating group membership). A significant odds ratio of 1.5 for a predictor, for example, indicates that on average, the odds are 50 percent higher of recidivating for youths with a "1" for that variable. In contrast a significant odds ratio of 0.5, indicates that on average, the odds of recidivating for youths with a "1" for that variable are 50 percent lower. Age at first arrest and all five of the neighborhood-level variables were standardized before being entered into the models. As a result, their odds ratios are interpreted slightly differently. An odds ratio of 1.5 for one of these variables indicates that on average, a one standard deviation increase will result in a 50 percent increased odds that a juvenile will recidivate.

Each set of two models for the four outcome measures will be discussed separately. There is very little variance in the models measuring the same outcomes – the same predictors are found to be significant with only minor variance in the effect size of predictors. Odds ratios for predictors are listed first for the model that includes disadvantage, followed by crime. A comparison of the models for each outcome follows.

Table 4.3: Odds Ratios of Individual- and Neighborhood-Level Predictors of Juvenile Recidivism

Model	1	2	3	4	5	6	7	8
Recidivism Offense	All	All	Drug	Drug	Perso	Perso	Propert	Propert
Individual-level								
Demographics								
White	0.96	0.87	0.70*	0.62*	0.82	0.83	1.28	1.25
Hispanic	1.08	1.05	1.42**	1.38*	0.70**	0.69**	0.99	1.00
Age at 1st Arrest	0.96	0.96	1.14**	1.13*	0.87**	0.87**	0.91*	0.91*
Family								
Parental Drug Abuse	1.10	1.10	0.98	0.98	1.16	1.16	1.06	1.06
Parental Criminality	1.14*	1.15*	1.05	1.05	1.29*	1.28*	1.09	1.09
Aftercare Case	1.52*	1.51*	1.71**	1.70*	1.05	1.05	1.16	1.15
On Public Assistance	1.10	1.10	1.10	1.10	0.96	0.96	1.12	1.13
Current Offense								
Sex Offense	0.53*	0.53*	0.57*		0.61*	0.61*	0.55**	0.55**
Person Offense	1.02	1.02	0.91	0.91	1.21	1.21	1.08	1.08
Property Offense	1.24*	1.24*	1.01	1.01	1.04	1.04	1.62**	1.62**
Criminal History								
On Probation	1.21*	1.22*	0.90	0.90	1.25	1.25	1.21	1.21
Prior Drug Offense	1.48*	1.49*	2.88**	2.89*	0.76*	0.76*	0.73**	0.73**
Prior Person Offense	1.25*	1.24*	1.14	1.13	1.34**	1.33**	0.96	0.96
Prior Placement	1.52*	1.52*	1.16	1.17	1.23	1.23	1.49**	1.49**
Neighborhood-level								
Disadvantage	1.10*	--	1.15**	--	1.01	--	1.02	--
Crime	--	1.06*	--	1.09*	--	1.02	--	1.00

** p < .01, * p < .05

Models 1 & 2: All Recidivism

Of the fourteen individual-level variables entered into the models, eight are found to be significant predictors of juvenile recidivism. Neither race nor "age at first arrest" are significant predictors of recidivism when all offense types are combined. Juveniles designated as aftercare cases have a relatively high odds ratio (*OR*=1.52; 1.51, $p < .01$), indicating that juveniles on an aftercare status are more likely (odds are more than 50 percent higher) to reoffend than juveniles who were not. Conversely, juveniles who committed a property offense (*OR*=1.24; 1.24 , $p < .01$) are more likely (odds 24 percent higher) to recidivate than juveniles who do not commit property offenses. All of the variables representing criminal history have a significant and positive relationship with the likelihood of recidivating. Having committed a prior drug offense (*OR*=1.48; 1.49, $p < .01$) and having had a prior out-of-home placement (any placement, dependent or delinquent, at any point in the youth's life; *OR*=1.52; 1.52, $p < .01$) are particularly strong predictors of juvenile recidivism.

The neighborhood disorder scale (*OR*=1.09, $p < .01$) is positively correlated with recidivism, indicating that as the level of disorder in a juvenile's neighborhood increases, so too does the odds that they will recidivate, while holding all other indicators constant. A similar relationship is found for neighborhood crime (OR = 1.06, p < .05), further indicative of the effects of neighborhood context on juvenile recidivism.

Models 3 & 4: Drug Recidivism

Models 2 and 3 indicate that six individual-level predictors are significantly correlated with juveniles who commit drug crimes as their recidivating offense. The effects of being Hispanic (*OR*=1.42; 1.38 , $p < .01$) and of being older at the time of first arrest (*OR*=1.14; 1.13, $p < .01$) are both positively related to drug recidivism. In contrast, white juveniles, on average, are less likely than black youths to recidivate through the commission of a drug crime (*OR*=0.70; 0.62, $p < .05$), holding other variables in the model constant. Juveniles on aftercare status are much more likely to reoffend (*OR*=1.71; 1.70, $p < .01$) than their non-placed counterparts. Of the current offense variables representing the offenses that brought the juveniles into our database, only sex offense is significantly related to drug recidivism (*OR*=0.57; 0.57, $p < .05$), and this relationship was negative. Only one of the criminal history predictors is significant for drug recidivism. Not surprisingly, prior drug offense (*OR*=2.88; 2.89, $p < .01$) exerts the greatest influence of all the variables in any of the models in this study. The effect of having a prior drug offense in a juvenile's criminal history is to nearly triple the odds that the juvenile will reoffend via a drug crime.

At the neighborhood-level, as in Models 1 and 2, both disadvantage (OR=1.15, p < .01) and crime (OR=1.09, p < .05) are significant predictors of juvenile drug recidivism.

Models 5 & 6: Person Recidivism

Models 5 and 6 possess two significant demographic variables. The first, Hispanic (*OR*=0.70; 0.69, $p < .01$), indicates that Hispanic juveniles are less likely to recidivate with a person offense than black youths. Age at first arrest has a positive effect (*OR*=0.87; 0.87, $p < .01$), signifying that juveniles who are older at the time of their first offense are similarly less likely to recidivate with a person offense. Only parent criminality is significant among the family context variables (*OR*=1.29; 1.28, $p < .05$): parental criminality is positively associated with person re-offending. Sex offense is a significant predictor as well (*OR*=0.61; 0.61, $p < .05$), reducing the likelihood of recidivating via a person offense for juveniles previously convicted of sexual offenses. Two criminal history variables exert significant influence on person recidivism, but in different ways. Having committed a prior drug crime significantly reduces the likelihood that a juvenile will recidivate with a person offense (*OR*=0.76; 0.76, $p < .05$). In contrast, having committed a prior person offense increases the likelihood of reoffending with a similar person offense (*OR*=1.34; 1.34, $p < .01$).

Neither neighborhood disadvantage nor crime is a significant predictor of person recidivism. This is largely due to a lack of significant variance in person offense recidivism between neighborhoods.

Models 7 & 8: Property Recidivism

As with Models 1 and 2, neither race variable is a significant predictor of property offense recidivism. Age at first arrest (*OR*=0.91; 0.91, $p < .05$) is negatively correlated, indicating the older a youth is, the less likely they are to recidivate with a property crime. No family context variables are significant predictors of property recidivism. Sex offense (*OR*=0.55; 0.55, $p < .01$), as with all other models, is negatively correlated with property recidivism. The effect of having committed a property offense as the current offense is to increase the odds of reoffending with a property offense by more than 60 percent (*OR*=1.62; 1.62, $p < .01$) over non-property offenders. Prior drug offenders are less likely to reoffend with a property offense (*OR*=0.73; 0.73, $p < .01$), while juveniles who have been placed out of their home in the past are nearly 50 percent more likely to recidivate with a property offense (*OR*=1.49; 1.49, $p < .01$) than are juveniles who have not been placed out of their homes.

Similar to person offense recidivism in Models 5 and 6, the neighborhood context attributes are not significantly correlated with property offense recidivism.

Discussion

The way in which we have parceled out the types of recidivating offense has allowed us to examine the individual- and neighborhood-level correlates of juvenile recidivism and estimate how their effects differ for drug, person, and property recidivating offenses. This research is among the first to ask this question and has uncovered several interesting findings that can be used to inform future research and juvenile justice practitioners. The first of these findings is that neighborhood context, in the form of disadvantage and crime, is a significant predictor of juvenile recidivism when offense type is ignored and when examining only drug offense recidivism.[i] These findings are important and consistent with the few studies of the neighborhood correlates of both adult and juvenile recidivism conducted in the past (Kubrin, Squires, & Stewart, 2007; Kubrin & Stewart, 2006; LeBaron, 2002; Simmons, 2001).

It is interesting to note that neighborhood context does not influence all types of recidivism offenses, but this is not surprising when considering that both race (Massey & Denton, 1989; Wilkes & Iceland, 2004) and recidivism offense type were found to be spatially clustered in Philadelphia after conducting preliminary hotspot analyses. Philadelphia has historically been, and is currently, a hypersegregated city (Massey & Denton, 1989; Wilkes & Iceland, 2004). If both recidivism offense type and race are highly segregated in Philadelphia, then perhaps race and neighborhood effects are confounding one another. This could help to explain why a neighborhood-level predictor representing racial heterogeneity was not found to significantly predict juvenile recidivism, and thus, was omitted from the final models. It may be that the racial and cultural effects created by the hypersegregated ethnic groups in Philadelphia are at least partially responsible for masking the more robust neighborhood effects that were hypothesized to be found. In contrast, the results of our individual-level analysis suggest that race is one of the strongest predictors of the type of recidivism offense. Race, at least to some extent, may also represent space in our study.

The significant effects of neighborhood-level variables that represent community processes are supportive of similar processes that are described by social disorganization theory. Our neighborhood disadvantage predictor can be interpreted to represent the poverty dimension of social disorganization theory and is further supportive of social disorganization theory by acting to increase rates of juvenile recidivism as disadvantage increases. The effects of our neighborhood crime predictor can be interpreted similarly, as it likely increases levels of disorganization within neighborhoods as it increases.

In comparing the effects of the individual-level variables across the models, it is first interesting to note how the effects of race differ by type of recidivism. In Models 1 and 2, we observe that being white or Hispanic, as compared to being black, is not predictive of recidivism. This lack of significance is soon explained by the opposing influences these two variables exert on specific types of recidivism. Being white decreases the likelihood of committing a drug recidivism offense, but increases the likelihood of committing a property crime as a recidivating offense. On the other hand, Hispanic juveniles are more likely than both blacks and whites to commit a drug crime as their recidivating offense, but less likely to commit a person offense. Age at first arrest also has varying effects across the models, with a positive correlation for drug recidivism, but a negative correlation with both person and property recidivism. These results can be interpreted to say that juveniles who offend earlier in their lives are more likely to recidivate with person and property offenses, while juvenile offenders who begin their criminal careers later are more likely to reoffend with drug crimes.

Of the family context predictors, parental criminality and being separated from the family either as a delinquent (aftercare status) or earlier (prior out-of-home placement) were significant predictors of recidivism. Surprisingly, when recidivism offense type is specified, parental criminality predicts person offenses, aftercare status predicts drug offending, and prior placement predicts property offending.

Current sex offense is the only indicator with a significant odds ratio less than 1.0; this finding is supported by the literature that describes the recidivism rates of juvenile sex offenders to be consistently lower than non-sex offender recidivism rates (Miner, 2002; Parks & Bard, 2006). Current person offense did not significantly predict any type of recidivism, but current property offense was positively correlated with property recidivism and all recidivism.

The most interesting and strongest individual-level predictors were derived from the juveniles' criminal history. Juvenile offense specialization, as mentioned above, is not commonplace. Few offenders specialize at all, and those individuals who are considered specialists generally age toward specialization (Farrington, Snyder, & Finnegan, 1988; Piquero, Paternoster, Mazerolle, Brame, & Dean, 1999). The results of this study, however, would appear to support the specialization of juvenile drug offenders. Juveniles with a prior drug offense were significantly more likely to re-offend with a drug crime, and were significantly less likely to re-offend with a person or property crime. More than any other type of crime studied in this analysis, prior drug offenses appear to lead to other drug offenses exclusively. This finding lends support to Chaiken and Johnson's (1988) study of types of juvenile drug offenders, in which they report the attribute of persistence amongst juvenile drug dealers. On the other hand, our findings do not support their characterization of juvenile drug dealers as being prone to violent offending.

Two-Level Analysis: Individuals and Programs

In the second HLM analysis, the neighborhood level variables were replaced by program level variables. These data were taken from the Program Design Inventory (PDI), a database of program design created to accompany the ProDES data.

Program-Level Variables

Data were collected from each of the programs regarding their structure, staffing, type and amount of services provided per week, and specific program goals. These data were reduced to create predictors that represent the structure, activity dosages, and level of cognitive-behavioral focus for each of the programs. The eleven program-level predictors used in this analysis are shown at the bottom of Table 4.4. The predictors representing program structure indicate that the programs, on average, have the capacity for 74 clients, have a curriculum lasting about 9.5 months, have four clients for every staff member, and have fewer than one licensed social worker.

Table 4.4: Descriptive Statistics of Individual- and Program-Level Variables

	Metric	N	Min	Max	M	SD
Individual-Level Predictors						
Demographics:						
White	0 = no, 1 = yes	7061	0	1	0.11	0.32
Hispanic	0 = no, 1 = yes	7061	0	1	0.13	0.34
Age at First Arrest	Continuous	7061	9	19	14.20	1.69
Family:						
Parental Drug Abuse	0 = no, 1 = yes	7061	0	1	0.20	0.40
Parental Criminality	0 = no, 1 = yes	7061	0	1	0.16	0.36
Aftercare Case	0 = no, 1 = yes	7061	0	1	0.35	0.48
On Public Assistance	0 = no, 1 = yes	7061	0	1	0.31	0.46
Current Offense:						
Sex Offense	0 = no, 1 = yes	7061	0	1	0.06	0.24
Person Offense	0 = no, 1 = yes	7061	0	1	0.36	0.48
Property Offense	0 = no, 1 = yes	7061	0	1	0.32	0.46
Criminal History:						
On Probation	0 = no, 1 = yes	7061	0	1	0.10	0.30
Prior Drug Offense	0 = no, 1 = yes	7061	0	1	0.33	0.47
Prior Person Offense	0 = no, 1 = yes	7061	0	1	0.20	0.40
Prior Placement	0 = no, 1 = yes	7061	0	1	0.07	0.25
Program-Level Predictors						
Program Structure:						
Client Capacity	Continuous	26	10	300	74.15	64.19
Average Length-of-Stay	Continuous	26	3	24	9.48	5.90
Staff-to-Client-Ratio	Continuous	26	0.04	1	0.25	0.25
Licensed Social Workers	Continuous	26	0	3	0.77	0.86
Service Dosages:						
Individual Counseling	Continuous	26	0	5.5	1.50	1.66
Family Counseling	Continuous	26	0	2	0.29	0.57
Group Counseling	Continuous	26	0	10	3.10	3.12
Vocational/Job Training	Continuous	26	0	21	2.54	5.24
Substance Abuse	Continuous	26	0	5	0.92	1.60
Goals:						
Cognitive/Thinking	Continuous	26	0	4	0.64	0.91

Service dosage items were created by, in the cases of substance abuse and vocational/job training, combining the weekly dosages of activities related to those types of services provided by the programs. In the cases of individual counseling, family counseling, and group counseling, these items were taken directly from the dataset and also represent the hourly total of these services per week. Table 4.4 shows that family counseling is the least common service provided by the programs (average of 0.29 hours per week), whereas vocational/job training-related services (average of 5.24 hours per week) are the most common.

Lastly, an indicator termed "Cognitive/Thinking" was created to represent the extent to which programs have a goal of changing the ways in which their clients think. Each program listed a series of goals during the data collection process. A comparison of the program services and the listed goals, however, indicated that the goal-related items could not be used to accurately describe the programs. For example, programs that professed to have several goals related to reducing substance abuse among their clients had little or no substance abuse-related program services that were provided to their clients. As a result, most of these items were not included in the analysis. Exceptions were made for the six items used to create the predictor representing the extent to which programs claimed to modify the ways in which their clients think and behave. They are:

Admission of crime/problem

Develop decision-making skills

Increase self-awareness/self-understanding

Recognition of motivating factors for negative behaviors

Resolve underlying dysfunctional core beliefs

Accept responsibility for behaviors/actions

For each of these goals, programs indicated whether or not they possessed these goals (0 = no, 1 = yes). The resulting predictor combines the number of these goals for each program. Table 4.4 indicates that the maximum number of these goals for any program is four, while the average number of these goals for the programs is just below one (0.91).

Analysis Plan

In order to estimate the effects of the above-mentioned program-level predictors on juvenile recidivism while controlling for individual-level characteristics, hierarchical generalized linear models are used. A random intercept model is used to predict all four of the outcome measures.

Results

Results of the hierarchical models are shown in Table 4.5.

Table 4.5: Odds Ratios of Individual- and Program-Level Predictors of Juvenile Recidivism

	Model 1 All Recidivism	Model 2 Drug Recidivism	Model 3 Person Recidivism	Model 4 Property Recidivism
Individual-Level Predictors				
Demographics:				
White	0.86*	0.60**	0.83	1.20
Hispanic	1.18	1.47**	0.73*	1.07
Age at First Arrest	0.93	1.13*	0.86**	0.88**
Family:				
Parental Drug Abuse	1.07	0.94	1.17*	1.01
Parental Criminality	1.14**	1.10	1.27**	1.05
Aftercare Case	1.15*	1.27*	0.99	1.02
On Public Assistance	1.09	1.11	0.96	1.09
Current Offense:				
Sex Offense	0.61**	0.70	0.59*	0.54**
Person Offense	0.99	0.87	1.27	1.03
Property Offense	1.18*	0.96	1.07	1.55**
Criminal History:				
On Probation	1.21	0.91	1.27	1.20
Prior Drug Offense	1.48**	2.82**	0.76*	0.75**
Prior Person Offense	1.24**	1.14	1.36**	0.95
Prior Placement	1.38*	1.17	1.10	1.42*
Program-Level Predictors				
Program Structure:				
Client Capacity	1.25**	1.28	1.11	1.21*
Average Length-of-Stay	0.94	0.91	0.82	1.08
Staff-to-Client-Ratio	1.00	0.91	0.85*	0.94
Licensed Social Workers	0.98	0.97	1.11	0.95
Service Dosages:				
Individual Counseling	1.05	1.10	1.10	1.07
Family Counseling	1.00	0.93	1.02	1.00
Group Counseling	0.92	0.93	0.87*	0.97
Vocational/Job Training	1.13	1.14	1.11	1.07
Substance Abuse	0.96	1.06	0.75*	0.98
Goals:				
Cognitive/Thinking	0.96	0.82	1.03	0.92

Model 1 shows the effects of individual- and program=level predictors on recidivism. Odds ratios of the demographic variables indicate that age at first arrest is unrelated to re-offending (*OR* = 0.93, $p > .05$), but being white (*OR* = 0.65, $p < .05$) is associated with a reduced likelihood of recidivism, relative to black youths. Parental criminality (*OR* = 1.14, $p < .01$) and having been in a residential facility immediately prior to their initial offense (*OR* = 1.15, $p < .05$) both increase the odds of a new offense. A prior out-of-home placement (*OR* = 1.38, $p < .05$) similarly increases the likelihood of re-offending. Model 1 indicates that prior drug offenses (*OR* = 1.48, $p < .01$) and prior person offenses (*OR* = 1.24, $p < .01$) exert a significant influence on increasing recidivism. In contrast, juvenile offenders in our sample who committed a sex offense are much less likely to recidivate (*OR* = 0.61, $p < .01$). One program-level predictor significantly influences the likelihood that juvenile offenders will recidivate: client capacity (*OR* = 1.25, $p < .01$). This indicates that clients at larger programs are more likely to reoffend.

Models 2-4 provide comparisons of predictors for each of the three recidivism offenses types. Model 2 is especially interesting, as drug offenders appear to be very different when compared to person and property reoffenders. Drug recidivists are more likely to be older at the time of their first arrest (*OR* = 1.13, $p < .05$) and Hispanic (*OR* = 1.47, $p < .01$) than other reoffending youths in the sample. Moreover, they seem to be more likely to specialize; evidenced by the significant value for the prior drug offense predictor (*OR* = 2.82, $p < .01$) representing an almost 300% increase in the likelihood of recidivating with a drug offense for juveniles with a prior drug offense. Further separating drug recidivists from their fellow offenders, is a lack of treatment correlates to influence their likelihood of reoffending: none of the ten program-level variables have a significant effect on drug recidivism.

Model 3 shows that person offense recidivists are less likely to be white (*OR* = 0.83, $p > .05$) or Hispanic (*OR* = 0.73, $p < .05$), and to come from families characterized by family-related issues, such as parental criminality (*OR* = 1.27, $p < .01$) and parental drug abuse (*OR* = 1.17, $p < .05$). Prior person offenses increase the likelihood of recidivating with a person offense (*OR* = 1.36, $p < .01$), while a prior drug offense has a significant, but opposite effect (*OR* = 0.76, $p < .01$), which further supports the specialization hypothesis. Person offense recidivists are the subpopulation of reoffenders that are most acutely affected by program-level predictors. Three program variables predict the likelihood of person recidivism: staff-to-client ratio (*OR* = 0.85, $p < .05$), group counseling dosage (*OR* = 0.87, $p < .05$), and the dosage of services aimed at reducing substance abuse among treatment clients (*OR* = 0.75, $p < .05$).

Model 4 shows that, compared to other reoffenders, juveniles who recidivate with a property offense are younger (*OR* = 0.88, $p < .01$) and not more or less likely to be of a specific race. Similar to person recidivists, property reoffenders are unlikely to have ever committed a prior drug offense (*OR* = 0.75, $p < .01$), but are likely to have committed property offense as their initial offense (*OR* = 1.55, $p < .01$), again suggesting the specialization of juvenile offenders. The effects of treatment on property recidivism mirror those of the aggregate recidivism measure: only the client capacity (*OR* = 1.21, $p < .05$) of programs exerts a significant influence on the likelihood that juveniles will recidivate with a property offense.

Discussion

This phase of the analysis indicates that program-level correlates of juvenile recidivism have differing effects based on recidivism offense type. It seems clear from these models that client capacity, or program size, increases levels of recidivism. But perhaps the most interesting finding is that program characteristics do not seem to have any effect on drug recidivists. While this may be initially surprising, it is important to remember that drug recidivists were the only youths in the neighborhood-level analysis initially described in this chapter that were impacted by neighborhood context. Clearly, drug recidivists are different.

These youths are likely to be embedded in environments in which drug offending is supported and encouraged. The relatively large effect of prior drug offending suggests offense specialization. Given that aftercare status (having reentered the community after a period of institutional commitment) predicts drug reoffending, we would expect that the concerns raised by Dishion, McCord, and Poulin (1999) regarding delinquent peer contagion may be influencing this relationship, and perhaps even contributing to the inability of program components to influence the likelihood of drug reoffending that is concluded from the HLGM results in Table 4.5.

Person offenders, in contrast, seem to be coming from families that are in disarray, in which parents are poor role models and may even be modeling the behavior that got these youths in trouble with the law. Such individual-level relationships with person recidivism are supportive of the relationships found between substance abuse services, group counseling, and staff-to-client ratio that were detected at the program-level. The finding that counseling and substance abuse services have positive effects on juveniles who are more likely to have criminal and drug-abusing parents is not hard to believe. The correlation between staff-to-client ration and person recidivism is at first counterintuitive, as a higher staff-to-client ratio (fewer staff members per client) is found to increase the likelihood of reoffending, but this relationship may be tied to the effects of group counseling. Group counseling is shown to work for person recidivists, and it may be that programs providing group counseling for its clients may naturally have higher staff-to-client ratios as a result of the nature of putting clients into group-based treatment.

Cross-Classified Model

In the final HLM analysis, both neighborhood- and program-level predictors were included in the models at level-two to create cross-classified HLM models. As each juvenile offender in the analysis is nested in both a program and a neighborhood, this technique is appropriate. More importantly, it will allow us to simultaneously estimate the effects of all three levels of predictors included in the earlier HLM analyses.

Program-Level Variables

The program-level predictors included in this analysis were also included in the previously described two-level model that examined program-level effects of recidivism. See the description of those indicators in that section. Descriptive statistics of these variables are shown again in Table 4.6.

Neighborhood-Level Variables

As in the two-level neighborhood analysis, a scale representing neighborhood disadvantage has been constructed in order to control for neighborhood context in this analysis. What differs from the earlier analysis is that this indicator, and the neighborhood-level of aggregation, are based on the 45 neighborhoods defined by the Philadelphia Health Management Corporation (PHMC). Intended to capture distinct communities in Philadelphia, the PHMC has identified 45 neighborhoods within the city. This change in neighborhood level of aggregation stems from the nature of cross-classified models. The data requirements of cross-classified HGLM models suggest the utilization of fewer neighborhood-level units in the current analysis.

Data aggregated to the 45 neighborhoods defined by the PHMC come from the PHMC's Southeastern Pennsylvania Household Heath Survey, a biannual survey of residents within five counties in Southeast Pennsylvania, regarding issues related to health and community. Data from surveys given in 2000, 2002, and 2004 were averaged, so that neighborhood change from 2000 to 2004 could be accounted for. In order to ensure that data from the survey respondents comprised a representative sample of Philadelphia residents, the survey data were weighted, per year, based on Census data from the 2000 Census Public Use Microdata Sample (PUMS) file. PUMS data contain information from one resident per Census household and were data cross-tabulated by demographic indicators and compared to the same indicators for PHMC survey respondents. Weight values for each year of the PHMC survey were calculated based on comparisons between residents in the two files (see Garcia et al., 2007 for a more about the weighting procedure).

As a result, the construction of a scale representing neighborhood disadvantage differs slightly from the census-based scale in the two-level analysis. The number and proportion of female headed-households is not available in the PHMC data, and is therefore not included in the current operationalization of neighborhood disadvantage. In its place, neighborhood household income (transformed to represent its inverse), was included due to its correlation with the other four items: neighborhood proportions on welfare, on public assistance, unemployed, and residents who are black. During the factor analysis process, the proportion of black residents was found to weaken the subsequent scale item, and was not included. The resultant neighborhood disadvantage predictor exhibits a high level of internal reliability (Cronbach's alpha = .947).

Table 4.6: Descriptive Statistics of Individual-, Neighborhood-, and Program-Level Variables

Individual-Level	Metric	N	Min	Max	M	SD
Demographics:						
White	0 = no, 1 = yes	7061	0	1	0.11	0.32
Hispanic	0 = no, 1 = yes	7061	0	1	0.13	0.34
Age at First Arrest	Continuous	7061	9	19	14.20	1.69
Family:						
Parental Drug Abuse	0 = no, 1 = yes	7061	0	1	0.20	0.40
Parental Criminality	0 = no, 1 = yes	7061	0	1	0.16	0.36
On Public Assistance	0 = no, 1 = yes	7061	0	1	0.31	0.46
Current Offense:						
Aftercare Case	0 = no, 1 = yes	7061	0	1	0.35	0.48
Sex Offense	0 = no, 1 = yes	7061	0	1	0.06	0.24
Person Offense	0 = no, 1 = yes	7061	0	1	0.36	0.48
Property Offense	0 = no, 1 = yes	7061	0	1	0.32	0.46
Criminal History:						
On Probation	0 = no, 1 = yes	7061	0	1	0.10	0.30
Prior Drug Offense	0 = no, 1 = yes	7061	0	1	0.33	0.47
Prior Person Offense	0 = no, 1 = yes	7061	0	1	0.20	0.40
Prior Placement	0 = no, 1 = yes	7061	0	1	0.07	0.25
Program-Level						
Program Structure:						
Client Capacity	Continuous	26	10	300	74.15	64.19
Average Length-of-Stay	Continuous	26	3	24	9.48	5.90

Staff-to-Client-Ratio	Continuous	26	0.04	1	0.25	0.25
Licensed Social Workers	Continuous	26	0	3	0.77	0.86
Service Dosages:						
Individual Counseling	Continuous	26	0	5.5	1.50	1.66
Family Counseling	Continuous	26	0	2	0.29	0.57
Group Counseling	Continuous	26	0	10	3.10	3.12
Vocational/Job Training	Continuous	26	0	21	2.54	5.24
Substance Abuse	Continuous	26	0	5	0.92	1.60
Goals:						
Cognitive/Thinking	Continuous	26	0	4	0.64	0.91
Neighborhood-Level						
Disadvantage*	Scale Item	45	-1.50	1.89	0	1.0
Poverty	Proportion	45	0.05	0.53	0.22	0.12
Public Assistance	Proportion	45	0	0.18	0.58	0.48
Unemployment	Proportion	45	0	0.17	0.78	0.45
Income	Continuous	45	-12.08	-4.9	-8.16	1.74

*Cronbach's α = 0.947

Results

Results of the cross-classified models are shown in Table 4.7. The results largely support the findings of the earlier two-level HLM analyses. Regarding neighborhood-level effects, the results from the cross-classified models are similar in that neighborhood disadvantage is a significant predictor of drug recidivism, but not person or property recidivism.

The effects of the program-level predictors are similarly supportive of the earlier two-level analysis, although there are some differences now that neighborhood context is also being controlled for in these models. Only client capacity significantly influences the aggregate recidivism measure, as in the two-level model, and does so in the same and expected direction. Model 2, estimating effects on drug recidivism, demonstrates the greatest change from the earlier two-level model with the inclusion of neighborhood disadvantage. Now, client capacity and the predictor measuring the level of cognitive goals are significant predictors of drug recidivism. Both are in the expected direction, as an increase in the capacity of a program increase the likelihood of recidivism, and an increase in the number of cognitive goals decreases the likelihood of recidivating with a drug crime.

The significant relationship found between cognitive goals and drug recidivism is interesting, but hardly surprising. Research has identified a lack of interpersonal skills among juvenile drug users that make them particularly responsive to cognitive-behavioral treatment (Catalano et al., 1991; Robertson, 2000, Veysey, 2008). Regarding juvenile drug sellers specifically, Schreiber (1992) found that juvenile drug sellers, when compared to other juvenile offenders, exhibited issues related to immaturity, decision-making, and thinking processes that are precisely what cognitive-behavioral therapy seeks to improve. What is surprising is that the cognitive goals predictor is not associated with the other outcomes, as numerous studies of juvenile offender treatment have identified cognitive-behavioral therapy as one of the most effective types of treatment (Guerra et al., 2008; Izzo and Ross, 1990; Landenberger and Lipsey, 2005; Lipsey, 1995, 1999).

Three factors are likely to push youths into drug selling: opportunities to engage in drug selling, pressure from peers to sell drugs and alienation from conventional social expectations. Opportunities to engage in drug trafficking are greater in neighborhoods where there is weak social organization and little in the way of social controls (Sampson & Groves, 1989). Some of this opportunity is real and some is perceived. Stanton and Galbraith (1994) argue this point well, reporting that youths who sell drugs are more likely to believe that adults and other youths in their neighborhood are also selling drugs. These perceptions become part of the rationale a youth uses to excuse involvement in disapproved behavior.

Table 4.7: Odds Ratios of Individual-, Neighborhood, and Program-Level Predictors of Juvenile Recidivism

	Model 1 All Recidivism	Model 2 Drug	Model 3 Person	Model 4 Property
Individual-Level				
Demographics:				
White	0.89	0.66**	0.82	1.18
Hispanic	1.10	1.32*	0.70*	1.09
Age at First Arrest	0.94*	1.13**	0.86**	0.88**
Family:				
Parental Drug Abuse	1.08	0.93	1.17	1.01
Parental Criminality	1.13	1.09	1.27*	1.05
Aftercare Case	1.20*	1.29*	0.99	1.02
On Public Assistance	1.08	1.08	0.96	1.10
Current Offense:				
Sex Offense	0.56**	0.70	0.59*	0.54*
Person Offense	1.00	0.87	1.26	1.03
Property Offense	1.21*	0.97	1.07	1.55**
Criminal History:				
On Probation	1.21*	0.91	1.26	1.20
Prior Drug Offense	1.49**	2.76**	0.77*	0.75*
Prior Person Offense	1.24**	1.14	1.35**	0.95
Prior Placement	1.39**	1.16	1.10	1.42*
Neighborhood-Level				
Disadvantage	1.09*	1.24**	1.00	0.97
Program-Level				
Program Structure:				
Client Capacity	1.27**	1.29**	1.11	1.22*
Average Length-of-Stay	0.95	0.90	0.82	1.07
Staff-to-Client-Ratio	0.97	0.90	0.85	0.94
Licensed Social Workers	0.98	0.99	1.11	0.95
Service Dosages:				
Individual Counseling	1.08	1.10	1.10	1.07
Family Counseling	0.98	0.91	1.02	1.01
Group Counseling	0.91*	0.93	0.87*	0.97
Vocational/Job Training	1.08	1.16	1.12	1.07
Substance Abuse	0.91	1.04	0.75**	0.98
Goals:				
Cognitive/Thinking	0.90	0.81*	1.03	0.91

Little and Steinberg (2006) found that alienation from conventional values and goals, as well as a loss of commitment to school were associated with involvement in drug selling. Similar findings of alienation are noted by Centers and Weist (1998), and Dembo et al. (1998). This distancing from values and goals of the larger society and the neighborhoods that promote them, leaves a void that can be filled by more proximate value messages. That the pressure to sell drugs can come from peers or family members has been documented by Stanton and Galbraith (1994) as a key predictor of drug selling. Little and Steinberg (2006) found, however, that youths who demonstrated autonomy from peer influence were more likely to sell drugs other than marijuana, suggesting that personal goals rather than external pressure drive decisions to sell drugs. In fact, they argue, financial rewards are more important than social rewards in their decisions to sell drugs.

There is an obvious financial gain to be made from selling drugs. If youths perceive their opportunities to earn decent incomes through socially acceptable means to be cut off and find that selling drugs earns them a respectable wage, then the rational choice is to sell drugs (Fagan & Freeman 1999; Nagin &Paternoster 1993; Stanton & Galbraith, 1994; Blumstein, 1993). In other words, decisions to sell drugs involve rational choice. But as Baumer and Gustafson (2007) argue well, using an anomie theory perspective, the combination of a commitment to monetary success goals, combined with economic disadvantage produces incentives to commit instrumental crimes such as drug selling. With neighborhood disadvantage predicting drug selling but not person or property reoffending, specialization in drug offending and cognitive program goals combining to predict drug reoffending, we have strong indication that an anomic perspective is valuable to understanding the spatial (and ethnic) concentrations of drug selling. These forces that influence drug dealing are clearly part of the cognitive landscape enabling youths to choose to sell drugs. We would expect that effective intervention will have to address those perception, values, and beliefs that underlie decisions to engage in drug selling. It may also be necessary to address neighborhood-level perceptions of the causes of economic disadvantage.

The results of Model 3 are very similar to the results of Model 3 in the earlier two-level analysis of program effects. As with the earlier analysis, the number of hours per week of group counseling and substance abuse services are significant predictors of person recidivism. This model deviates from the earlier analysis because staff-to-client ratio is no longer a significant predictor.

The literature on peer contagion suggests that grouping youths together for purposes of treatment undermines the treatment process. High concentrations of delinquent youths, exclusive of prosocial adolescents, create opportunities for deviancy training. According to Dishion, Dodge and Lansford (2006), some factors likely to mitigate this iatrogenic effect include limiting the amount of time delinquent youths are together relative to their mixing with youths outside the group, providing structured group activities that minimize unsupervised interaction, creating small, diverse groups of youths that are unlikely to be friends outside of the group setting, and locating the program in an area where delinquent youths outside the program are unlikely to congregate. These suggestions are consistent with the effectiveness of group treatment when the program capacity is small, the program serves the entire city, and the program’s dosage measures are high, meaning that the time spent in the program is devoted to treatment activities. That is, small community-based programs that provide high levels of supervised group activities to youths from different neighborhoods are less likely to produce high levels of deviancy training than large programs with significant amounts of unstructured time, particularly when program size creates opportunities for youths to gather who are already acquainted.

Lastly, Model 4 describes results that are identical to the earlier two-level analyses of property recidivism. Only client capacity, among program-level predictors, exerts a significant influence on property recidivism, and in same and expected direction.

The individual-level effects are strongly supportive of both of the earlier two-level analyses. As with those analyses, there are stark differences observed between the individual-level predictors of drug recidivism and the other outcomes. The demographic predictors are perhaps the most interesting, as they indicate that older juveniles and Hispanics are more likely to recidivate with a drug crime than any other offense type. These findings are in contrast to those in Model 3 that indicate the opposite effects for age and Hispanics on person recidivism.

CHAPTER 5
A FOCUS ON DRUG OFFENDING WITH GEOGRAPHICALLY WEIGHTED REGRESSION

INTRODUCTION

We have seen from several of our analyses, reported in earlier chapters of this report, that drug offenses differ from other offenses in two important respects: they are more predictable and they are more spatially dependent. Geographically weighted regression (GWR), then, is an appropriate tool for exploring further the relationship between our predictors of interest and juvenile recidivism. GWR adds to our prior analyses by generating parameters that are disaggregated by the spatial units of analysis that we have selected.

GEOGRAPHICALLY WEIGHTED REGRESSION (GWR) RESULTS

This analysis maps the results of geographically weighted regression (GWR), a statistical tool used to determine spatial nonstationarity, in order to investigate the causes of juvenile drug-crime recidivism. Specifically, this portion of the study investigates how the influence of individual and socioeconomic characteristics on juvenile drug-crime recidivism varies across neighborhoods in Philadelphia.

Through the course of this study, drug-crimes and drug-crime recidivism have been demonstrated to be a different type of crime than personal, property and weapons offenses. Therefore, we found it desirable to devote a portion of this study solely on drug offenses. In the GWR analysis we concentrated on recidivating offenses involving drugs.

The geocoded home addresses of 7,171 juvenile delinquents were integrated with both individual- and neighborhood-level data that were determined to be significant during the HLM analysis. These variables included: public assistance income, any prior charges for sexual offense, number of prior delinquency offences, age, White (dummy) variable, Hispanic (dummy) variable, on probation, parental drug- and crime-activity, the number of times the juvenile lived in an institutional setting, the number of times the youth was in an out-of-home placement and the type of his original offense (drug, personal offense or property offense). Descriptive statistics are presented in Table 5.1.

Table 5.1
Descriptive Statistics

Variable	Number	Percent	Minimum	Maximum	Mean
Race					
African American	5257	73.3			
Hispanic	943	13.1			
Caucasian	818	11.4			
Asian	122	1.7			
Other	31	< 0.4			
Age			10	20	15.7
Any Charges for Sexual Offense	427	6			
On Probation	704	9.8			
Parental Characteristics					
Neither	4129				
Drug abuse	1464	20.4			
Criminality	1149	16			
Both	429	5.9			
Instant Offense					
Person crime	2571	35.8			
Property crime	2284	31.8			
Drug crime	1691	23.6			
Other	625	8.7			
Recidivating Offense	2890	40.3			
Drug crime	1030	14.4			
Person crime	726	10.1			
Property crime	796	11.1			
Weapons crime	147	2			
Other crimes	191	2.6			
Number of Prior Delinquency Arrests			0	14	
0	3928	54.8			
1	1787	24.9			
2	802	11.1			
3	352	4.9			
4	172	2.4			
5 or more	130	1.8			
Public assistance	2271	31.7			
Institution	2556	35.7			

Prior out of home placement	546	7.6

Results of conventional forward-stepwise logistic regression are reported in Table 5.2. The model suggest that significant variables include the juveniles' age, race, poverty, instant drug offenses or prior sex offenses and institutionalization influence the likelihood of recidivism with a drug crime.

Table 5.2

	Linear Regression Model
ROC Curve	0.690**
Constant	0.017***(114.261)
Nagelkerke (R^2)	0.096
Correct (%)	85.6
Count (n)	7171
Classification Accuracy	0.14
Individual	
Public Assistance	-
Sex Offense	0.431***(12.825)
Prior Delinquency Offenses	1.093***(12.332)
Age	1.119***(21.027
White	0.552***(18.429)
Hispanic	1.654***(31.372)
Probation	-
Parent Drug	-
Parental Crime	-
Live Institution	1.589***(39.006)
Prior Out of Home Placement	-
Instant Drug	2.277***(122.557)
Instant Person	-
Instant Property	-

* Significance < 0.05

** Significance < 0.01

*** Significance < 0.001

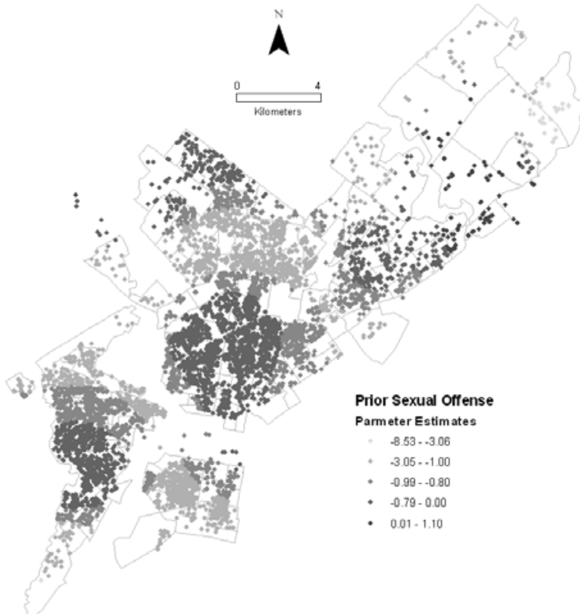
Values indicate the odds ratio. Wald statistic is reported in parentheses.

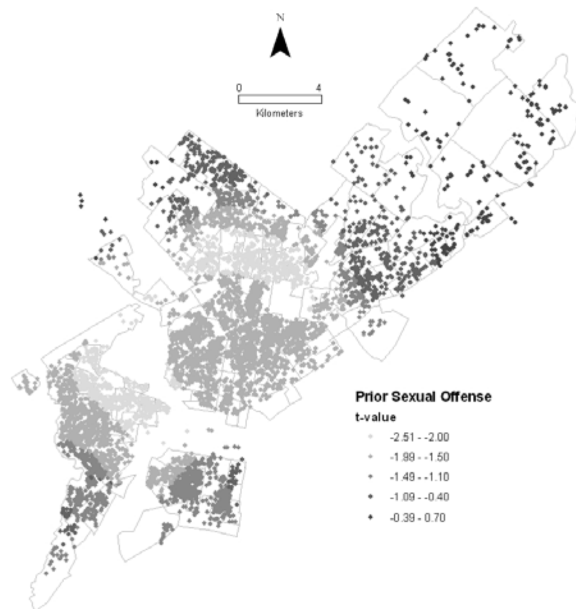
Notice, however, that the model is poorly specified. It only explains 9.6% of the variation in drug-crime recidivism because this global, or average, estimate acts as a false-constant across the region. Using the GWR results, the analysis will be able to attain a high degree of specificity as it relays the data in a localized manner.

The seven significant variables from the logistic regression were entered into a geographically weighted logistic regression and the results were mapped for visual analysis of spatial nonstationarity.

The choice of bandwidth is an important setting in GWR as it specifies the radius over which a relationship is assumed to exist among observations. The bandwidth may be set manually to a fixed radius or may be adaptive such that the radius is allowed to vary according to the spatial density of samples (in our case, juveniles) – the bandwidth expands in regions of sparse sampling and contracts in regions of dense sampling. We experimented with a variety of bandwidth settings, including fixed and adaptive settings. Because the density of juveniles in our study varies widely, generally according to the distribution of population density, we initially experimented with an adaptive bandwidth. However, the results suggested that the bandwidth setting was to coarse in many regions, producing overly smooth maps of the spatial variation in coefficient values that appeared to mask substantial local variation. We ultimately settled on a fixed bandwidth setting of 2 km, which allows enough observations in sparsely sampled areas to generate meaningful parameter estimates yet doesn't unduly mask local variation of estimates in densely sampled areas. Recall that the impact of a relatively high bandwidth in densely sampled regions is offset by the distance decay function of the GWR itself, which gives greater weight to observations nearer to the observation on which the current GWR iteration is focused.

Maps 1-8 suggest that the magnitude and significance of the explanatory variables' influence on drug recidivism varies across neighborhoods. Individual points on the maps indicate home addresses of the 7,171 juvenile delinquent males in the study. Map pairs for each variable include parameter estimate map and a t-value map, thus allowing a visually examination of spatial nonstationarity.



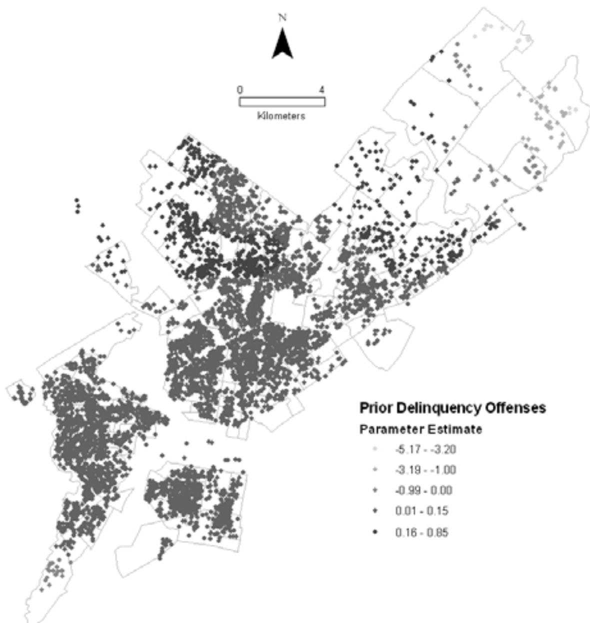


Map 5.1.a

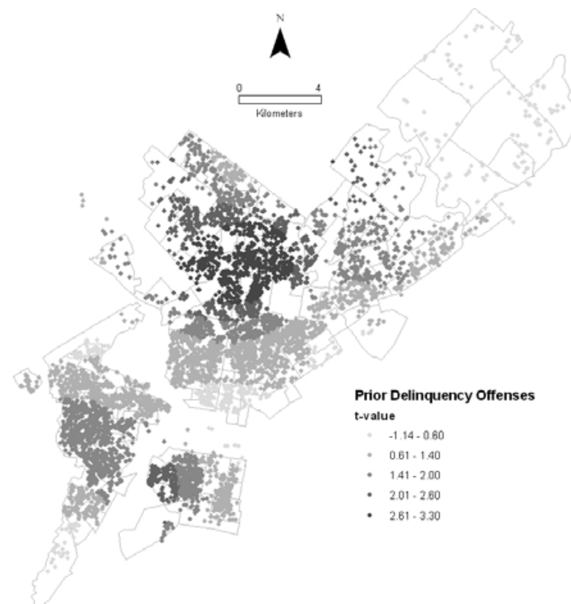
Map 5.1.b

According to our logistic regression model, which provides a global parameter estimate, youths who have prior sexual offense charges are not likely to reoffend with a drug crime. However, according to the map of GWR results, the relationship between prior sexual offense and recidivating drug offense is not stationary throughout the city. In some areas of the city, a strong negative relationship exists; while in others, a weak positive relationship is present.

Map 5.1.b, the t-value map, shows areas where there is a significant relationship between the explanatory and the dependent variables. Significant relationships are exclusively negative, and occur inparts of West Philadelphia, stretching from Mantua west to Wynnefield, as well as in parts of northern Philadelphia, stretching from East Falls east to Olney.

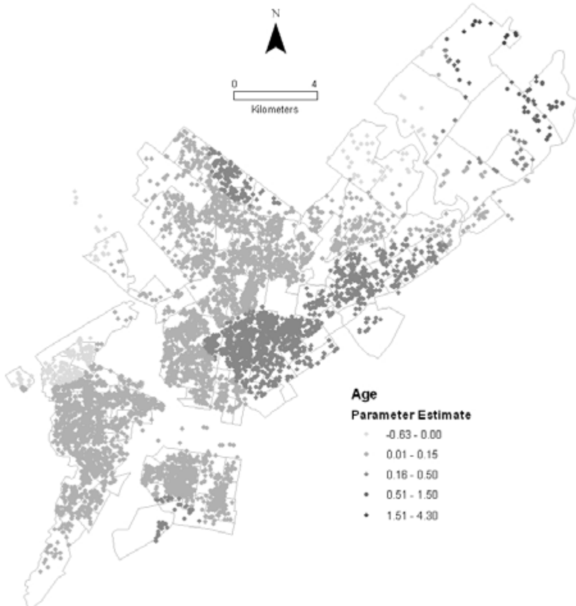


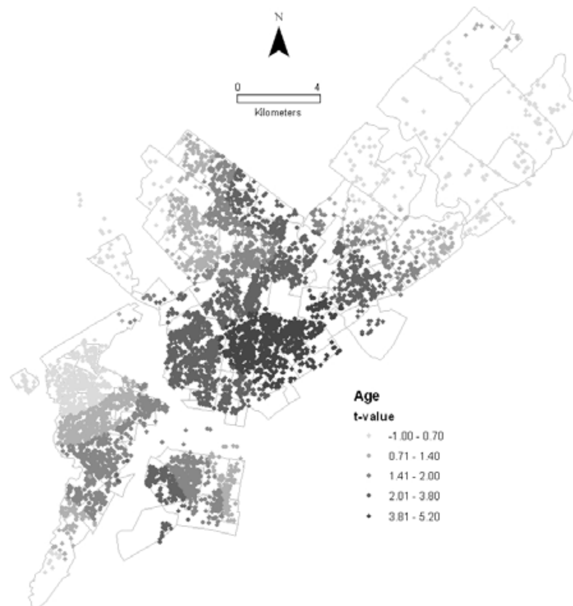
Map 5.2.a



Map 5.2.b

Even though the **number of prior delinquency** offenses does not seem to have a strong correlation in our global estimates, the GWR analysis demonstrates that this relationship is not stationary throughout Philadelphia. For example, in the less densely populated Northeast, there is a strong negative correlation; with every additional past offense, the less likely it is that a youth will recidivate with a drug crime. However, the t-value map in Figure 5.2.b shows that the Northeast has a low significance level. The t-value map demonstrates that values in East Falls are significant correlation values are weak.

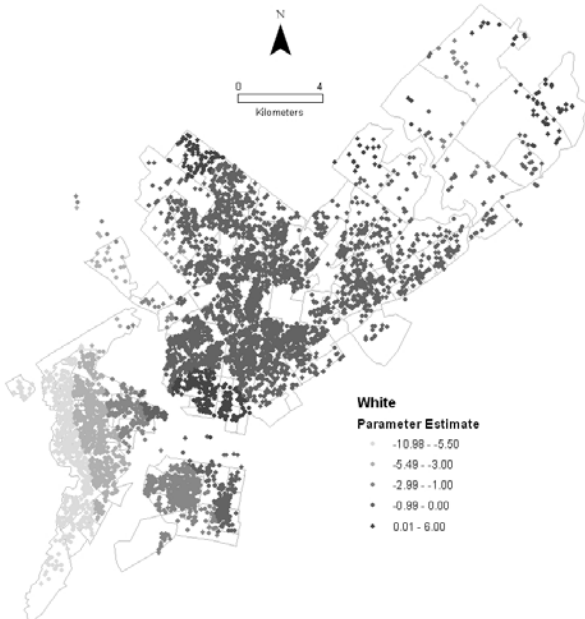


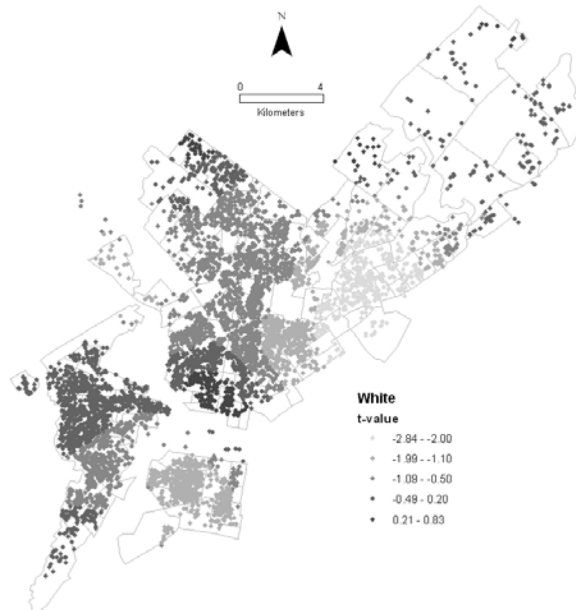


Map 5.3.a

Map 5.3.b

According to the logistic regression and GWR analysis, **age** is positively related to drug-crime recidivism; with each year increase in age, the increased likelihood of drug recidivism increases. However, in examining the GWR results in Map 5.3.b, one can see that in some areas of the city, particularly the Hunting Park neighborhood, these parameter estimates are strongly positive.

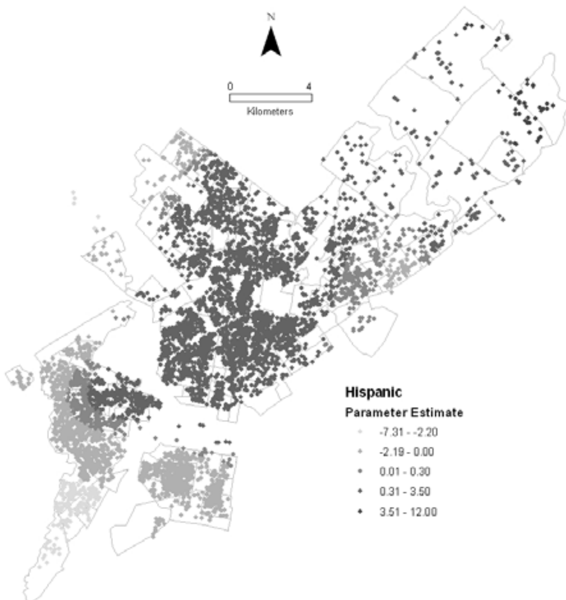


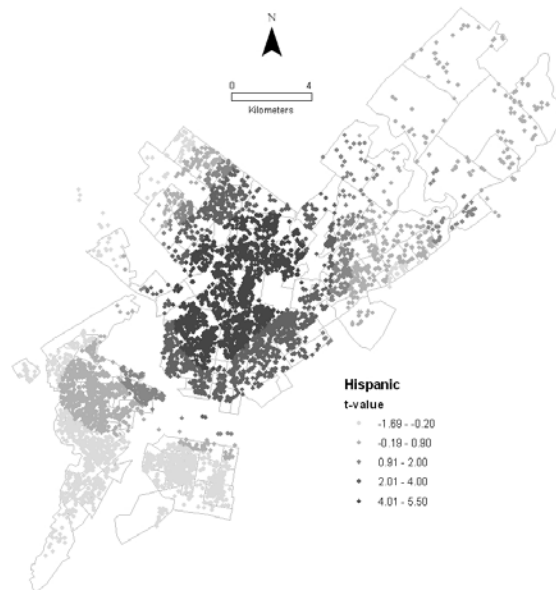


Map 5.4.a

Map 5.4.b

If a Philadelphia youth is **White**, he is less likely to reoffend with a drug crime, according to the global estimate. However, the GWR reveals that this relationship is not constant throughout the city. From the t-value map (5.4.b), we see that in the area around Frankford, the relationship is negative and statistically significant.



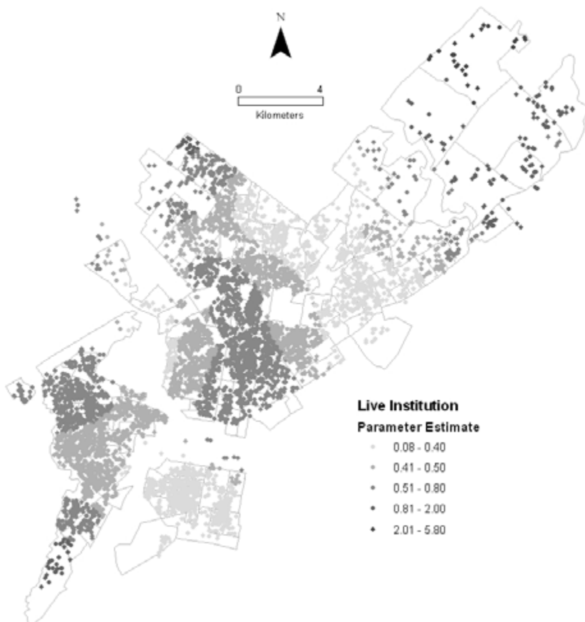


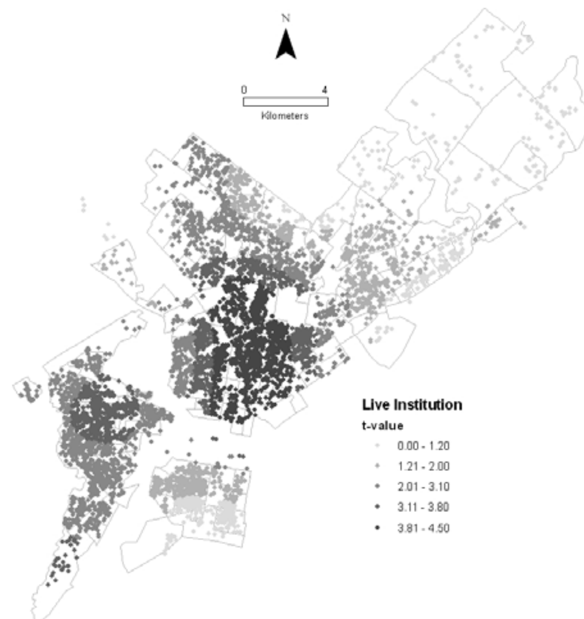
Map 5.5.a

Map 5.5.b

In stark contrast to Map 5.4.b, youths who self-identify as **Hispanic** are more likely to recidivate with a drug crime in most areas of the city. This can be observed particularly in areas of Hispanic concentration, as in Hunting Park, and in African American neighborhoods in North Philadelphia that border Hunting Park (to the north, west, and south). This pattern is likely due to the GWR picking up on the strong relationship between race and drug crime recidivism, which can be observed most clearly in areas within the bandwidth that straddle different ethnic neighborhoods, or for neighborhoods that include substantial numbers of residents of different ethnicities. It is interesting to note that the significance increases in the positive direction over areas with both African American and Hispanic residents, and decreases (even becoming negative, though not significant) in areas under the bandwidth that contain mainly white and Hispanic residents. This suggests that the propensity of Hispanics to recidivate with a drug crime is based on the reduced likelihood of African Americans to recidivate with a drug crime in, and nearby, Hispanic neighborhoods. In those same neighborhoods, Hispanic youths are not more likely than white youths to recidivate with a drug crime.

One can speculate that with limited economic opportunities due to citizenship concerns, Hispanic youths may become involved in drug-crime activities to supplement other sources of income.

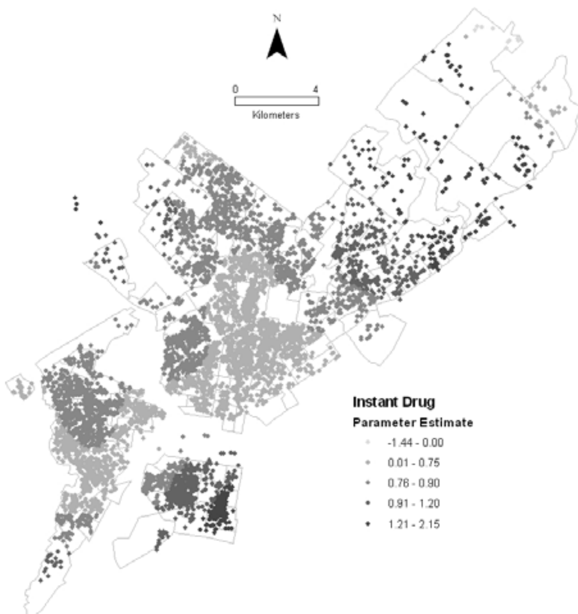




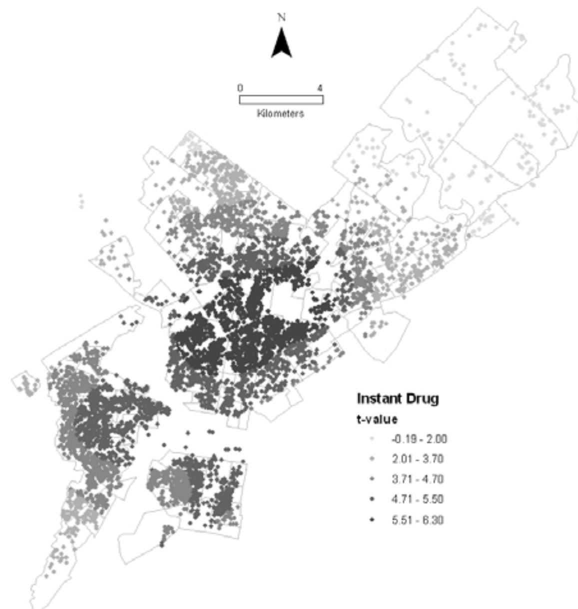
Map 5.6.a

Map 5.6.b

The localized view of the variable relaying the number of times a juvenile delinquent was placed in institutional care immediately prior to their current community-based experience (**aftercare cases**) reveals that this variable is significant and positive primarily in portions of northern Philadelphia and West Philadelphia, primarily disadvantaged African American and Hispanic neighborhoods. It is interesting to note that the highest parameter estimates occur in the far Northeast, though the t-values indicate the relationship is not significant. It is somewhat odd that this area contains both the highest parameter estimates and the lowest t-values in the city, but this is also likely due to the lower sampling density in the northeast, which can lead to high standard errors and thus low t-values, even when the magnitude of the effect is substantial. One might speculate that in this area the effect of the institutional placement is very inconsistent, but substantial for certain juveniles who have been placed in institutional care. Perhaps this has to with peer contagion in the context of institutional placement, where those juveniles residing in the Northeast, which is predominantly white and middle class and has a low rate of juvenile delinquency and drug recidivism generally, mix with residents of socioeconomically distressed neighborhoods, where delinquency and drug recidivism are substantially more prevalent.



Map 5.7.a



Map 5.7.b

Juveniles, whose **drug offense** was the crime that placed them in the data set, are likely to reoffend with a drug crime. Beyond a few individuals in the far Northeast, this relationship is statistically significant throughout the city in the local logistic regression analysis. This relationship is consistent with our conventional knowledge of drug crimes requiring relationships with others who deal in drugs; a youth who knows where, how and from whom to obtain drugs is more likely to use that knowledge and those relationships again to commit additional drug offenses.

The mapped GWR results spotlight the cluster in South Philadelphia where there is a strong and statistically significant relationship between instant drug crime offenses and drug crime recidivating. This visually-recognized hotspot is near the intersection of Interstate-95, Interstate-76, the Walt Whitman Bridge to New Jersey and the official sports stadiums of Philadelphia. One could speculate that this transportation and activity hub provides quick access to a variety of individuals, therefore the means for a successful drug crime location.

The GWR analysis suggests that substantial spatial nonstationarity exists in the data set with regards to the explanatory variables associated with drug recidivism. These results add credence to the idea that subgroups of juveniles may exhibit similar qualities and causal mechanisms regarding recidivism, and therefore that different explanatory models of recidivism should be developed for these subgroups. In GWR, as opposed to PLAID or other methods of identifying subgroups, space serves as the organizing framework for dividing the data set into subgroups of juveniles. Here, variation in causal mechanism over space can be interpreted as a proxy for variation among groups of similarly behaving juveniles. It is likely that the nonstationarity exhibited in the data set may be partially captured by space (or the tendency for similarly behaving kids to live nearby one another) and partly by conditions that are generally non-spatial (i.e. kids with family members who have criminal backgrounds). Thus, one of the major challenges for this research in the future is to identify approaches to defining optimal sub-groups to maximize predictive power based on both spatial and nonspatial characteristics.

CHAPTER 6
SPECIALIZATION AND PEER CONTAGION

INTRODUCTION

Forces found to influence juvenile recidivism rates include individual, family, program, and neighborhood-level factors. And while knowledge of the impact of environmental forces on delinquency has been developing for some time (Beyers, Loeber, Wikstrom, and Stouthamer-Loeber, 2001; Bursik, 1988; Liberman, 2007; Loeber and Wikstrom, 1993; Sampson and Groves, 1989; Sampson, Morenoff, and Earls, 1999; Simcha-Fagan and Schwartz, 1986; Hawkins, Catalano and Arthur, 2002), relatively little attention has been given to the environmental factors that increase the likelihood of recidivism. Moreover, as the results reported in Chapter 4 indicate, there is reason to pursue the concepts of peer contagion and specialization, and to do so using a spatial perspective.

DEVIANT PEER CONTAGION

The general thesis of peer contagion in the context of juvenile delinquency is expressed by Dodge et al. (2006: 3-4) as "[d]eviant adolescents become more deviant by associating freely with deviant peers." In fact, they observe, "deviant peer affiliation is a stronger predictor of delinquent behavior than such variables as family, school and community characteristics." In the present paper we argue that high concentrations of delinquent youths are a community characteristic that produces higher rates of recidivism. The impact of deviant peer contagion has been the focus of recent research on the potential harm of aggregating delinquent youths for purposes of rehabilitation (Dishion and Dodge, 2006)). The concept of deviant peer contagion derives from several sources and can be traced to concerns present at the inception of the juvenile court (Bernard, 1992; Tanenhaus, 2004). Founders of the juvenile court discussed evidence of the deleterious effects of committing children to correctional facilities that also housed more experienced juvenile and adult criminals.

A similar concern underlay the emergence of deinstitutionalization arguments during the 1970's, during which time an anti-institutionalization theme emerged that recognized the harmful effects of delinquent subcultures that emerge in reaction to low levels of opportunity to accumulate social capital (e.g. Cohen, 1955), as well as the stigma attached to being labeled and treated as a delinquent (Lemert, 1967). Environmental forces and government decisions were seen as having the effect of pushing youths toward increasing involvement with delinquent peer groups.

Importantly for our purposes, concentrations of delinquent peers can also mean isolation from prosocial peers. This concern was raised by Tannenbaum (1938) in the first expression of labeling theory. In both neighborhoods and schools, when the ratio of delinquents (especially repeat offenders) to nondelinquents reaches a certain tipping point, delinquent youths may have little access to nondelinquents youths but many opportunities to interact with delinquent youths. This imbalance of peer group orientation increases chances for delinquent peer influence to dominate.

Another source of knowledge regarding deviant peer contagion is program evaluation research. Dishion et al. (1999) demonstrated that programs can create harm when they create groups of exclusively delinquent youths. The process of contagion is expressed as "deviance-training" that includes identity formation, communication of antisocial norms, sensitivity to peers who have similar life experiences, and direct recruitment into gangs or deviant groups. Follow-up studies of youth attending the Cambridge-Somerville Youth Study Evaluation of the 1940's showed similar iatrogenic, or treatment-induced, negative effects (McCord, 2003). This early study used a matched sample of control and treatment boys to evaluate program effects. When these youth were followed up into middle age via court records, vital statistics, mental health records, and treatment centers, it appeared as though the treatment program had negative effects (McCord, 2003). McCord (2003) found that boys who received the most treatment over the longest period of time were most likely to have iatrogenic effects.

The process by which contagion occurs is referred to as "deviancy training" (Dishion & Dodge, 2006, p. 28). The general notion is that the motivation to commit specific types of offenses is learned through interactions with peers. This view is related to theories of differential association (Short and Strodbeck, 1965) and, more specifically to reinforcement theory (Burgess and Akers, 1966). Both theories rely heavily on social learning perspectives (Bandura, 1969). Both the behavior and the motivation to commit deviant acts are learned through interactions with others, and these others are most likely to be peers.

Specialization in Crime Type

Offense specialization has been the subject of research for both adults and juveniles and has implications for both theory and responses to crime. If offending is specialized, it may be necessary to formulate different explanations of offending for different types of offenses. If there is a high level of diversity in offending, then offense type does not inform the development or testing of theory. There is some indication that offense-specific prediction can produce greater explanatory power than general models, at least with respect to sex offending (Lussier, LeBlanc, and Proulx, 2005), but most research on juveniles emphasizes generality in offending (Bursik, 1980; Cohen, 1986; Klein, 1984; Piquero, Paternoster, Mazerolle, Brame, and Dean, 1999). Klein (1984) observed that although there is a general assumption that adolescents are diverse in their offending, there is frequent reference to types of offenders, using an offense label such as auto thief, to characterize individuals. His own survey of the literature leads him to conclude, however, that versatility is the norm.

Improved measurement, has, however presented new evidence that specialization may occur among juvenile offenders. This new research has focused on testing the relative impact of individual propensity as compared to opportunity, the impact of age on specialization, and the impact of age of delinquency onset on specialization. Tests of the relative impact of age and age of onset have led to the conclusion that age, not age of onset, increases the likelihood that offenders will specialize (Piquero, Paternoster, Maxerolle, Brame, and Dean, 1999). Bursik (1980) also found some evidence of racial differentiation in specialization, but there is little emphasis on this variable since then.

More recent research has found, however, that the pattern of specialization differs by type of offense. Armstrong (2008) tested if both the impact of age and extent of arrest influence the likelihood of specialization. He found that age does affect specialization for property and violent offending, but that for drug and miscellaneous offending age had little impact. Specialization in drug offending increased over arrests irrespective of age. This finding suggests that drug offending may differ from property and violent offending in that drug offending may be conditioned by differential opportunities provided by the neighborhoods in which offenders reside. With regard to person and property offenses, however, specialization increased across arrests but decreases in specialization were found when controls for age were introduced. Age, then, seems to affect person and property offense specialization but not drug offense specialization.

Several studies of delinquency have specified the type of delinquency that certain neighborhood features are more apt to influence. Jacob (2006) found that residential mobility is the best predictor of juvenile property crime while the rate of lone-parent families is the best predictor of violent crime. The work of Osgood and Chambers (2000) to test social disorganization in rural areas found that the percentage of female-headed households in a neighborhood was the strongest predictor of violent crime committed by juveniles. Sampson and Grove's (1989) test of social disorganization in Great Britain found that their constructs of organizational participation and local friendship groups were the strongest predictors of burglary, while ethnic heterogeneity significantly predicted only property crime (which included vandalism that Sampson and Groves consider to be a crime of juveniles). "Family disruption" was found to predict violent crime and the measure of "unsupervised peer groups" was found to be predictive of both property and violent crime (Sampson and Groves, 1989). These findings are supportive of an investigation of the impact of social disorganization on juvenile recidivism and, further, of an investigation that distinguishes juvenile recidivism by offense type.

DATA

We repeat here information that is also presented in Chapter 1. Data on juvenile delinquents were acquired from the ProDES (Program Development and Evaluation System) database, developed by the Crime and Justice Research Institute at Temple University under a contract with the City of Philadelphia. The ProDES database tracks juveniles assigned to court-ordered programs by the Family Court of Philadelphia, Pennsylvania and was designed to evaluate all programs used by the City of Philadelphia for its delinquent youths. ProDES is case-based, with a 7-year sample of 26,464 cases (10,980 youths) with cases in family court between 1996 and 2003. A case was defined as the period from any new commitment to a program for a new offense or probation violation, or change in program commitment (including movement from residential care to aftercare), to six months following discharge from the program.

ProDES collected data at four points in time: (1) at the point of disposition (the juvenile equivalent of sentencing), data were culled from the youth's record that contains information such as offense history, placement history, needs (e.g., drug use, mental health problems) and family history; (2) at program intake, staff persons were asked to complete a needs assessment and the youth completes a self-report section containing psychometric scales; (3) at discharge, the intake process was repeated and program staff reported on the youth's progress in the program; and (4) six months following program discharge, a follow-up record check was conducted to identify any new petitions (arrests leading to charges) generated in the juvenile or adult court systems, and telephone interviews were conducted with youths, when available, and guardians. Although the sample ranges in age from 10 to 20 years old, the majority (69%) are between 15 and 17 years old. These cases are primarily male (90%) and African American (73%).

The cases in ProDES were geocoded based on the home address given at the point of commitment to a community-based program. We also restricted our analysis to youths which had been in the system at least six months, so as to examine only those cases which had the possibility of recidivating. Of those cases, we eliminated female cases from the analysis, as the literature, as well as our own preliminary analyses, suggested that the causes of female juvenile recidivism differ from those of male juvenile recidivism. We also eliminated from our analysis period of time in which a youth was placed in a residential program (which would thus render the environmental variables of the juvenile's residence location moot). Youths on aftercare were included, since these were youths who were attending community-based programs. Aftercare status was coded in order to account for the possibility that youths who had been confined immediately prior to their community program were reacting to that experience. These criteria result in 11,016 remaining cases in ProDES.

Note that a single juvenile may be listed as multiple cases within the database, if the juvenile continues to re-offend after completing a court-ordered program. An analysis at the case level may thus be biased by particular juveniles who habitually reoffend. To address this issue, we created a new data set by selecting only the first community-commitment case for each individual juvenile. This new juvenile-level data set encodes characteristics of the juvenile at the time he entered the Family Court system in addition to the characteristics of the next, i.e. recidivating offense, if one exists. This 'juvenile-level' data set contained 7,166 records. We considered several outcome variables based on different recidivating offense types: drug offenses, person (violent) offenses, and property offenses. We also considered a recidivism outcome variable of any offense type. Each of the four outcomes is dichotomous – whether the juvenile reoffended with an offense of that type. Of the 7,166 juveniles in the data set, 1,030 (14%) recidivated with a drug offense, 725 (10%) recidivated with a person offense, and 794 (11%) recidivated with a property offense. Recidivism with other offense types (e.g. sex offenses, weapons offenses) was less common. The number of juveniles recidivating with any offense type was 2,881 (40%).

Four types of explanatory variables are used in this study to address the theoretical mechanisms described in the literature review: 1) background characteristics of the individual juvenile, 2) the initial offense that the juvenile committed upon entry to the Family Court system (referred to as the 'instant offense'), 3) social disorganization within the neighborhood within which the juvenile resides, and 4) indicators of overall delinquency and recidivism nearby the juvenile's residence (referred to as 'contagion' variables). Tables 6.1 and 6.2 report descriptive statistics for all the categorical and continuous explanatory variables used in this study, respectively.

Table 6.1. Descriptive Statistics for Categorical Explanatory Variables (*N*=7,166)

Explanatory Variable	N	%
Race		
White	818	11%
African American	5,252	73%
Hispanic	943	13%
Other	153	3%
Public assistance	2,271	32%
Parental crime	1,149	16%
Prior institutional living arrangement	2,553	36%
Out-of-home placement ever	545	8%
Instant Drug Offense	1,691	24%
Instant Person Offense	2,571	36%
Instant Property Offense	2,280	32%

Table 6.2. Descriptive Statistics for Continuous Explanatory Variables (*N*=7,166)

Explanatory Variable	Mean	SD
Age (years)	15.7	1.7
Number of prior arrests	0.8	1.3
Drug sale density (per km^2)	214	316
% female HH w/ children	20%	8%
% vacant housing	15%	10%
% high school graduate	59%	14%
Area Juvenile count	183	100

Area any recidivism rate	0.40	0.68
Area drug recidivism rate	0.15	0.54
Area person recidivism rate	0.10	0.34
Area property recidivism rate	0.11	0.36

Variables describing background characteristics of the individual juvenile include basic descriptors of age and race. Also included is poverty status, which is captured by encoding whether the family received public assistance income (welfare). The juvenile's family history regarding crime was captured by a variable indicating whether a parent of the juvenile has a criminal record. The juvenile's own criminal history was captured using variables that indicated the number of prior arrests (note that a juvenile may have been previously arrested but not sent to a court-ordered program), whether the juvenile was living in an institution (as opposed to with his family or other living arrangement) immediately prior to the targeted community-based case, and whether the juvenile had any out-of-home placement prior at any time since birth. Note that the "lives in an institution" variable indicates a juvenile with severe enough delinquent behavior or other issues for a judge to decide that it was in the community's best interest to remove the juvenile from his home.

To investigate crime specialization, the instant offense for each juvenile included as an explanatory variables, coded in the same manner as the outcome variables: drug offense, person offense, or property offense. Social disorganization of the juvenile's residential neighborhood was captured using crime, housing, and socioeconomic data. Block-level addresses of arrest data for the period 2000-2002 were acquired from the Philadelphia Police Department as text addresses and then geocoded. We then calculated for each juvenile the density of arrests in their home neighborhood by summing the number of arrests within 500 meters of each juvenile's home and dividing the sum by the area of the circular neighborhood defined by that 500 meter radius. We focused on two types of police-data arrests for our study: drug sale arrests and person offense arrests, as both variables are indicative of social disorganization. Because the two variables are highly correlated, and we found in univariate tests that drug sale arrests had a stronger relationship with our outcomes than person offense arrests, we utilized drug sale arrests exclusively as a crime-based indicator of social disorganization.

We also considered a variety of housing and socioeconomic variables derived from U.S. Bureau of the Census 2000 block group level data, including metrics of race, poverty, public assistance, and other characteristics. After investigating correlations among these and other explanatory variables used in the study, we focused on three independent Census variables that are intended to capture family organization, housing infrastructure, and educational attainment in the neighborhood within which juveniles reside: the percent of households which are female-headed with children, the percent of housing units that are vacant, and the percent of the population over the age of 25 with a high school diploma or equivalent.

In considering the influence of peer-contagion, we hypothesized that the likelihood of a juvenile recidivating may be influenced by the behavior of juveniles living nearby. For example, the likelihood of a juvenile recidivating with a drug crime may be enhanced if other juvenile delinquents nearby are recidivating with drug crimes at a relatively high rate. We generated five contagion effect variables to capture this effect. The first ('area juvenile count') is simply the number of juvenile delinquents, as encoded within the data set, who live within one kilometer of the juvenile's home. The second variable ('area any recidivism rate') is the total recidivism rate for all the juvenile delinquents who live within one kilometer – i.e. the proportion of these nearby juveniles that recidivated (with any offense type). A similar procedure was applied to drug recidivating offenses, person recidivating offenses, and property recidivating offenses in order to generate 'area drug recidivism rate,' 'area person recidivism rate,' and 'area property recidivism rate' variables, respectively.

METHODS

The application of offense transition analysis to the measurement of specialization is dominated by the Forward Specialization Coefficient (FSC), which we employ here to capture the degree of crime specialization as represented by those juveniles in our data set who recidivate. Since the FSC is a forward oriented measure, it measures the extent to which offenders with a specific type of offense, say violent, also have a violent offense as their second offense. When the FSC score equals 1, we assume complete forward specialization, and when the score is 0, we assume complete versatility. Calculation of the FSC begins with an offense transition matrix consisting of the joint distribution of two consecutive offense types. If we consider a matrix of instant and recidivating offense types labeled types *j* and *k*, respectively, then

$$\mathrm{FSC}_{jk} = \frac{O_{jk} - E_{jk}}{R_{jk} - E_{jk}}$$

where *O* is the number of observed cases, *E* is the number of cases expected by chance alone, and *R* is the number of cases in the row. The FSC can also result in a negative score, meaning that the likelihood of a second offense of a specific type is reduced by its occurrence as the first offense (Farrington, Snyder, and Finnegan, 1988).

According to Paternoster et al. (1998), the FSC approximates a normal distribution, and thus in samples of sufficiently large size (as with the present study) the standard error, SE, can be estimated as

$$\mathrm{SE}_{FSC_{jk}} = \sqrt{\frac{C_k(N - R_j)}{R_j N(N - C_k)}}$$

where C is the column total and N is the total number of cases in the matrix. We employ the standard error to test whether the FSC for any instant and recidivating offense transition is significantly different from zero at the 95% confidence level.

We also explored univariate relationships of each of our explanatory variables with each of the outcome variables. For dichotomous variables, we employed the Chi-square test to test whether there was a significant difference in frequency of occurrence between those juveniles who recidivated (or recidivated with a particular type of offense) and those who did not. For continuous variables we employed the Mann-Whitney U test to test whether there was a significant difference in the ranks of the means between juveniles who recidivated and those who did not.

For multivariate analysis, we employed stepwise-forward logistic regression to test the association of the explanatory variables with each of our four outcome variables. This approach is a well-established method for reducing the number of explanatory variables in a regression model by iteratively adding explanatory variables to the regression equation only when their relationship with the outcome is significant, after taking into account the influence of the other explanatory variables already present in the model (Darlington, 1990). Such an approach aids in the development of parsimonious models and interpretation of the regression. The stepwise procedure was carried out using four blocks of explanatory variables, where block 1 consisted of the variables describing the characteristics of the individual juvenile, block 2 entered the juvenile's instant offense type, block 3 consisted of variables indicating the social disorganization of the juvenile's residential neighborhood, and block 4 consisted of the spatial contagion variables. If an explanatory variable is entered into a model during the stepwise procedure in one block, it is kept in the models for the subsequent blocks.

Continuous variables were transformed by taking the natural log. We were careful to develop models without issues of multicollinearity. A few of our explanatory variables were collinear with Pearson correlations between 0.50-0.70, notably area juvenile count with area drug recidivism rate and percent female headed household. The stepwise approach assisted in addressing this issue – no models combined variables that are highly collinear with each other in a single logistic regression equation.

RESULTS

Offense Specialization

Results of calculating the FSC for each combination of instant and recidivating offense type indicates that having an instant offense of a particular offense type is associated with a recidivating offense of the same type (Table 6.3). This is especially true for recidivating drug offenders (FSC = 0.40). Those with an instant drug offense were also negatively associated with a recidivating offense of any other type, as one would expect.

Table 6.3. Forward Specialization Coefficient

	Recidivating Drug	Recidivating Person	Recidivating Property	Recidivating Other
Instant Drug	0.40*	-0.10*	-0.13*	-0.18*
Instant Person	0.02	0.11*	0.04*	-0.17*
Instant Property	0.04	-0.01	0.14*	-0.17*
Instant Other	0.13*	-0.03	0.05	-0.15*
* p<0.05				

Univariate Results

Results of the Chi-square tests are reported in Tables 4 and 5, where Table 4 reports the results for any recidivism and drug offense recidivism and Table 5 reports the results for person and property offense recidivism. Tables 6 and 7 report the results of the Mann-Whitney U tests, where Table 6 reports the results for any recidivism and drug offense recidivism and Table 7 reports the results for person and property offense recidivism.

Table 6.4. Chi-Square Tests of Any Recidivism and Drug Recidivism

Variable			Any Recidivism		Chi-Square	Drug Recidivism		Chi-Square
			0	1	0	1	0	
White	0	Obs. Exp.	3753 3796	2595 2552	10.55***	5382 5436	966 912	32.18***
	1	Obs. Exp.	532 489	286 329		754 700	64 118	
African American	0	Obs. Exp.	1152 1145	762 770	0.17	1616 1639	298 275	3.04
	1	Obs. Exp.	3133 3141	2119 2112		4520 4497	732 755	
Hispanic	0	Obs. Exp.	3764 3721	2459 2502	9.34***	5414 5329	809 895	
	1	Obs. Exp.	521 564	422 379		722 808	221 136	
Public Assistance	0	Obs. Exp.	3019 2927	1876 1968	22.68***	4243 4191	652 704	13.93***
	1	Obs. Exp.	1266 1358	1005 913		1893 1945	378 326	
Parental Crime	0	Obs. Exp.	3657 3598	2360 2419	15.04***	5169 5152	848 865	2.39
	1	Obs. Exp.	628 687	521 462		967 984	182 165	
Prior institutional living arrangement	0	Obs. Exp.	2991 2758	1622 1855	136.94***	4084 3950	529 663	88.84***
	1	Obs. Exp.	1294 1527	1259 1026		2052 2186	501 367	
Out-of-home placement ever	0	Obs. Exp.	4058 3959	2563 2662	80.78***	5710 5669	911 952	26.68***
	1	Obs. Exp.	227 326	318 219		426 467	119 78	
Instant Sex Offense	0	Obs. Exp.	3951 4030	2788 2709	64.11***	5729 5770	1010 967	34.64***
	1	Obs. Exp.	334 255	93 172		407 366	20 61	
Instant Drug Offense	0	Obs. Exp.	3371 3274	2104 2201	30.39***	4877 4688	598 787	224.52***
	1	Obs. Exp.	914 1011	777 680		1259 1448	432 243	
Instant Person Offense	0	Obs. Exp.	2616 2748	1979 1847	43.72***	3807 3935	788 661	80.17***

	1	Obs. Exp.	1669 1537	902 1034		2329 2202	242 370	
Instant Property Offense	0	Obs. Exp.	2972 2922	1914 1964	6.79**	4130 4184	756 702	15.08***
	1	Obs. Exp.	1313 1363	967 917		2006 1952	274 328	
"Obs." indicates the observed frequency. "Exp." indicates the expected frequency. *p<0.05, **p<0.01, ***p<0.005								

Table 6.5. Chi-Square Tests of Person Recidivism and Property Recidivism

Variable			Person Recidivism		Chi-Square	Property Recidivism		Chi-Square
			0	1	0	1	0	
White	0	Obs. Exp.	5698 5706	650 642	0.91	5664 5645	684 703	5.25*
	1	Obs. Exp.	743 735	75 83		708 727	110 91	
African American	0	Obs. Exp.	1759 1720	155 194	11.71***	1682 1702	232 212	2.87
	1	Obs. Exp.	4682 4721	570 531		4690 4670	562 582	
Hispanic	0	Obs. Exp.	5565 5593	658 630	10.84***	5527 5534	696 690	0.52
	1	Obs. Exp.	876 848	67 95		845 839	98 105	
Public Assistance	0	Obs. Exp.	4404 4400	491 495	0.13	4379 4353	516 542	4.55*
	1	Obs. Exp.	2037 2041	234 230		1993 2019	278 252	
Parental Crime	0	Obs. Exp.	5442 5408	575 609	12.99***	5363 5350	654 667	1.69
	1	Obs. Exp.	999 1033	150 116		1009 1022	140 127	
Prior nstitutional living arrangement	0	Obs. Exp.	4169 4146	444 467	3.45	4144 4102	469 511	10.96***
	1	Obs. Exp.	2272 2295	281 258		2228 2270	325 283	
Out-of-home placement ever	0	Obs. Exp.	5973 5951	648 670	10.44***	5915 5887	706 734	15.37***
	1	Obs. Exp.	468 490	77 55		457 485	88 60	
Instant Sex Offense	0	Obs. Exp.	6048 6057	691 682	2.32	5970 5992	769 747	12.58***

	1	Obs. Exp.	393 384	34 43		402 380	25 47	
Instant Drug Offense	0	Obs. Exp.	4888 4921	587 554	9.32***	4804 4868	671 607	32.55***
	1	Obs. Exp.	1553 1520	138 171		1568 1504	123 187	
Instant Person Offense	0	Obs. Exp.	4169 4130	426 465	10.09***	4056 4086	539 509	5.49*
	1	Obs. Exp.	2272 2311	299 260		2316 2286	255 285	
Instant Property Offense	0	Obs. Exp.	4395 4392	491 494	0.08	4440 4345	446 541	59.39***
	1	Obs. Exp.	2046 2049	234 231		1932 2027	348 253	

"Obs." indicates the observed frequency. "Exp." indicates the expected frequency.
*p<0.05, **p<0.01, ***p<0.005

Table 6.6. Mann-Whitney U Tests of Any Recidivism and Drug Recidivism

Variable	Any	Mean Rank	Mann-Whitney U	Drug	Mean Rank	Mann-Whitney U
Age	0 1	3492 3720	5779547***	0 1	3493 4121	2606203***
Number of prior arrests	0 1	3319 3977	5038850***	0 1	3502 4068	2661274***
Density of drug sales	0 1	3541 3646	5991316*	0 1	3555 3753	2985891**
% Female HH w/ children	0 1	3495 3716	5792636***	0 1	3524 3941	2792074***
% Vacant housing	0 1	3495 3715	5793677***	0 1	3532 3892	2842260***
% High school graduate	0 1	3693 3420	5702673***	0 1	3657 3148	2711772***
Area juvenile count	0 1	3484 3732	5744465***	0 1	3497 4096	2631838***
Area any recidivism rate	0 1	3372 3898	5267387***	0 1	3500 4081	2647110***
Area drug recidivism rate	0 1	3466 3758	5669120***	0 1	3464 4297	2424848***
Area person recidivism rate	0 1	3492 3719	5780902***	0 1	3596 3512	3086205
Area property	0	3535	5963261*	0	3627	2891999**

recidivism rate	1	3656		1	3323	*
*p<0.05, **p<0.01, ***p<0.005						

Table 5.7. Mann-Whitney U Test of Person Recidivism and Property Recidivism

Variable	Per- son	Mean Rank	Mann- Whitney U	Prop -erty	Mean Rank	Mann- Whitney U
Age	0 1	3605 3397	2199393**	0 1	3592 3513	2473334
Number of prior arrests	0 1	3552 3863	2132286***	0 1	3536 3966	2226286***
Density of drug sales	0 1	3577 3639	2294657	0 1	3590 3533	2489196
% Female HH w/ children	0 1	3573 3673	2269558	0 1	3593 3506	2468187
% Vacant housing	0 1	3579 3628	2302656	0 1	3584 3583	2529634
% High school graduate	0 1	3588 3547	2308588	0 1	3580 3615	2504920
Area juvenile count	0 1	3594 3492	2268634	0 1	3601 3439	2415207*
Area any recidivism rate	0 1	3553 3858	2136084***	0 1	3566 3727	2415941*
Area drug recidivism rate	0 1	3590 3526	2292966	0 1	3606 3402	2385402**
Area person recidivism rate	0 1	3508 4252	1850268***	0 1	3579 3623	2498152
Area property recidivism rate	0 1	3576 3648	2288465	0 1	3510 4170	2063642***
*p<0.05, **p<0.01, ***p<0.005						

The univariate results indicate that recidivists (of any offense type) tend to be older, non-white, and poor, with evidence of prior involvement with the justice system for themselves as well as their parents. Recidivists also tend to live in neighborhoods characterized by social disorganization, and high concentrations of juvenile delinquents and recidivists of all offense types living nearby. Race associations with certain types of offenses emerge when looking at specific recidivating offense types. Drug, person, and property offense recidivism is associated with being Hispanic, African American, white, respectively. There is also variability in the association of parental criminality, social disorganization, and spatial contagion with different types of offense recidivation.

Multivariate Results

The results of the stepwise-forward regression of any recidivism is presented in Table 6.8, where models 1-4 contain the results of adding variable blocks 1-4 to the regression equation, respectively. Model 1, which consists of only the individual variables, indicates that the likelihood of recidivating is enhanced by being Hispanic, poverty, and prior involvement with the justice system. Model 2 indicates that juveniles with instant sex offenses are less likely to reoffend while those with drug or property offenses are more likely. Hispanic becomes not significant when the instant offense variables are entered. Model 3 indicates that only one social disorganization variable adds explanatory power to the model – the likelihood of recidivism is enhanced with an increasing percentage of female headed households with children in the juvenile's home neighborhood. This social disorganization variable becomes not significant in model 4, however, when accounting for the influence of the recidivism rate (of any offense type) among juvenile delinquents living nearby.

Table 6.8. Stepwise-Forward Logistic Regression of Any Recidivism (All Recidivating Offenses) (*N*=7,166)

Explanatory Variables	Model 1	Model 2	Model 3	Model 4
Individual				
Age	-	-	-	-
White	0.80** (7.79) C.I. 0.68-0.94	0.80** (7.69) C.I. 0.68-0.94	0.88 (2.19) C.I. 0.74-1.04	0.88 (2.24) C.I. 0.74-1.04
Hispanic	1.18* (5.21) C.I. 1.02-1.36	1.12 (2.35) C.I. 0.97-1.29	1.08 (1.08) C.I. 0.93-1.25	0.98 (0.07) C.I. 0.84-1.14
Public assistance	1.16** (7.38) C.I. 1.04-1.28	1.14** (5.96) C.I. 1.03 – 1.27	1.12* (4.08) C.I. 1.00-1.24	1.09 (2.34) C.I. 0.98-1.21
Parental crime	1.15* (4.24) C.I. 1.01-1.31	1.18* (5.73) C.I. 1.03-1.34	1.17* (5.43) C.I. 1.03-1.34	1.17* (5.27) C.I. 1.02-1.33
Number of prior arrests	1.24*** (105.01) C.I. 1.19-1.29	1.23*** (93.42) C.I.1.18-1.28	1.23*** (93.66) C.I. 1.18-1.28	1.22*** (90.67) 1.17-1.27
Prior institutional living arrangement	1.53*** (66.51) C.I. 1.38-1.70	1.51*** (60.49) C.I. 1.36-1.67	1.51*** (61.58) C.I. 1.37-1.68	1.50 *** (58.99) C.I. 1.36-1.67
Out of home placement ever	-	-	-	-
Instant Offense Type				
Sex offense		0.52*** (27.97) C.I. 0.40-0.66	0.52*** (27.55) C.I. 0.40-0.66	0.52*** (27.69) C.I. 0.40-0.66
Drug offense		1.32*** (18.88) C.I. 1.17-1.50	1.31*** (17.27) C.I. 1.15-1.48	1.29*** (15.78) C.I. 1.14-1.47

Person offense		-	-	-
Property offense		1.20*** (9.53) C.I. 1.07-1.35	1.20*** (9.19) C.I. 1.07-1.34	1.20*** (9.01) C.I. 1.06-1.34
Neighborhood Social Disorg.				
Area drug sale density (ln)			-	-
Area % female HH w/ children (ln)			1.20*** (11.15) C.I. 1.08-1.33	1.01 (0.05) C.I. 0.91-1.13
Area % vacant housing (ln)			-	-
Area % high school graduate (ln)			-	-
Contagion Effects				
Area juvenile count (ln)				-
Area any recidivism rate (ln)				4.01*** (67.84) C.I. 2.88-5.59
Area drug recidivism rate (ln)				-
Area person recidivism rate (ln)				-
Area property recidivism rate (ln)				-
Constant	0.45*** (398.80)	0.42*** (295.15)	0.25*** (72.33)	1.49 (2.40)
Nagelkerke R2	0.05	0.07	0.07	0.08
Area under ROC curve				0.64*** C.I. 0.63-0.66
A gray box indicates a variable that was excluded from that model run. A dash indicates a variable that was allowed to enter that model but was not included by the stepwise procedure. Cell values indicate odds ratios. Wald statistic shown in parentheses. "C.I." indicates confidence interval at 95% confidence. *p<0.05, **p<0.01, ***p<0.005				

Table 6.9 reports analogous results to Table 8, but for drug offense recidivism only. Results or model 1 are similar to Table 8, with the exception that poverty and a parental criminal record are not associated with drug offense recidivism, though increasing age is. Model 2 indicates that juveniles with an instant drug offense are more likely to reoffend with a drug offense while juveniles with an instant sex offense are less likely. Models 3 and 4 indicate that while the

percentage of vacant housing adds explanatory power to the model – the likelihood of drug offense recidivism increases as the proportion of vacant housing increases – this social disorganization variable is not significant once the effects of the spatial contagion variables are added. Drug offense recidivism is more likely in the presence of a high rate of drug recidivism nearby.

Table 6.9. Stepwise-Forward Logistic Regression of Drug Offense Recidivism (*N* =7,166)

Explanatory Variables	Model 1	Model 2	Model 3	Model 4
Individual				
Age	1.17*** (46.37) C.I. 1.12-1.23	1.12*** (21.09) C.I. 1.07-1.18	1.12*** (22.15) C.I. 1.07-1.18	1.13*** (24.05) C.I. 1.08-1.19
White	0.51*** (24.74) C.I. 0.39-0.66	0.55*** (18.51) C.I. 0.42-0.72	0.61*** (12.34) C.I. 0.46-0.80	0.66*** (8.35) C.I. 0.50-0.88
Hispanic	1.90*** (53.80) C.I. 1.60-2.26	1.65*** (31.26) C.I. 1.39-1.97	1.65*** (31.24) C.I. 1.37-1.97	1.20 (3.35) C.I. 0.99-1.45
Public assistance	-	-	-	-
Parental crime	-	-	-	-
Number of prior arrests	1.10*** (15.45) C.I. 1.05-1.16	1.09*** (12.32) C.I. 1.04-1.15	1.09*** (11.18) C.I. 1.04-1.15	1.09*** (10.34) C.I. 1.03-1.14
Prior institutional living arrangement	1.52*** (33.15) C.I. 1.32-1.76	1.59*** (39.11) C.I. 1.38-1.84	1.59*** (38.50) C.I. 1.37-1.83	1.59*** (38.79) C.I. 1.38-1.84
Out of home placement ever	-	-	-	-
Instant Offense Type				
Sex offense		0.43*** (12.87) C.I. 0.27-0.68	0.43*** (12.88) C.I. 0.27-0.68	0.43*** (12.65) C.I. 0.27-0.69
Drug offense		2.27*** (122.15) C.I. 1.97-2.63	2.24*** (117.76) C.I. 1.94-2.60	2.11*** (97.95) C.I. 1.82-2.44
Person offense		-	-	-
Property offense		-	-	-
Neighborhood Social Disorg.				
Area drug sale density (ln)			-	-
Area % female HH w/ children (ln)			-	-

Area % vacant housing (ln)			1.17** (7.87) C.I. 1.05-1.31	0.98 (0.08) C.I. 0.87-1.11
Area % high school graduate (ln)			-	-
Contagion Effects				
Area juvenile count (ln)				-
Area any recidivism rate (ln)				-
Area drug recidivism rate (ln)				2.57*** (70.34) C.I. 2.06-3.21
Area person recidivism rate (ln)				-
Area property recidivism rate (ln)				-
Constant	0.01*** (156.87)	0.17*** (114.37)	0.01*** (119.47)	0.10*** (21.84)
Nagelkerke R2	0.06	0.10	0.10	0.12
Area under ROC curve				0.71*** C.I. 0.69-0.72
A gray box indicates a variable that was excluded from that model run. A dash indicates a variable that was allowed to enter that model but was not included by the stepwise procedure. Cell values indicate odds ratios. Wald statistic shown in parentheses. “C.I.” indicates confidence interval at 95% confidence. *p<0.05, **p<0.01, ***p<0.005				

The logistic regression of person offense recidivism is presented in Table 6.10. Juveniles committing person offense recidivism are more likely to be younger and less likely to be Hispanic. Model 2 indicates that juveniles with an instant sex offense are less likely to reoffend with a person offense and juveniles with an instant person offense are more likely. No neighborhood social disorganization variables are associated with person offense recidivism; however a high person offense recidivism rate nearby increases the likelihood of a juvenile recidivating with a person offense, while also diminishing the influence of being Hispanic.

Table 6.10. Stepwise-Forward Logistic Regression of Person Offense Recidivism (*N* =7,166)

Explanatory Variables	Model 1	Model 2	Model 3	Model 4
Individual				
Age	0.92*** (11.73) C.I. 0.88-0.97	0.93*** (8.98) C.I. 0.89-0.98	0.93*** (8.98) C.I. 0.89-0.98	0.93*** (8.66) C.I. 0.89-0.98

White	-	-	-	-
Hispanic	0.65*** (10.58) C.I. 0.50-0.84	0.66*** (9.65) C.I. 0.51-0.86	0.66*** (9.65) C.I. 0.51-0.86	0.74* (5.17) C.I. 0.56-0.96
Public assistance	-	-	-	-
Parental crime	1.37 *** (10.11) C.I. 1.13-1.66	1.37*** (9.94) C.I. 1.13-1.66	1.37*** (9.94) C.I. 1.13-1.66	1.37*** (9.80) C.I. 1.12-1.66
Number of prior arrests	1.13*** (19.61) C.I. 1.07-1.20	1.13*** (19.20) C.I.1.07-1.20	1.13*** (19.20) C.I.1.07-1.20	1.13*** (17.27) C.I. 1.06-1.19
Prior institutional living arrangement	-	-	-	-
Out of home placement ever	-	-	-	-
Instant Offense Type				
Sex offense		0.63* (5.89) C.I. 0.43-0.92	0.63* (5.89) C.I. 0.43-0.92	0.62* (6.11) C.I. 0.43-0.91
Drug offense		-	-	-
Person offense		1.31*** (10.13) C.I. 1.11-1.54	1.31*** (10.13) C.I. 1.11-1.54	1.30*** (9.44) C.I. 1.10-1.53
Property offense		-	-	-
Neighborhood Social Disorg.				
Area drug sale density (ln)			-	-
Area % female HH w/ children (ln)			-	-
Area % vacant housing (ln)			-	-
Area % high school graduate (ln)			-	-
Contagion Effects				
Area juvenile count (ln)				-
Area any recidivism rate (ln)				-
Area drug recidivism rate (ln)				-
Area person recidivism rate (ln)				3.07*** (74.65) C.I. 2.38-3.96

Area property recidivism rate (ln)				-
Constant	0.36** (7.76)	0.29*** (10.64)	0.29*** (10.64)	3.67** (7.28)
Nagelkerke R2	0.01	0.02	0.02	0.04
Area under ROC curve				0.63*** C.I. 0.61-0.65
A gray box indicates a variable that was excluded from that model run. A dash indicates a variable that was allowed to enter that model but was not included by the stepwise procedure. Cell values indicate odds ratios. Wald statistic shown in parentheses. "C.I." indicates confidence interval at 95% confidence. *p<0.05, **p<0.01, ***p<0.005				

Table 6.11 indicates that juveniles reoffending with a property offense are more likely to be white and have a higher number of prior arrests. Model 2 indicates that having an instant sex or drug offense reduces the likelihood of reoffending with a property offense while having an instant property offense increases the likelihood. None of the neighborhood disorganization variables was entered into the model. As with the models of the other outcome variables, a higher rate of property offense recidivism nearby increases the likelihood that a juvenile will recidivate with a property crime.

Table 6.11. Stepwise-Forward Logistic Regression of Property Offense Recidivism (*N* =7,166)

Explanatory Variables	Model 1	Model 2	Model 3	Model 4
Individual				
Age	-	-	-	-
White	1.27 * (4.68) C.I. 1.02-1.58	1.17 (1.87) C.I. 0.94-1.45	1.17 (1.87) C.I. 0.94-1.45	1.00 (0.00) C.I. 0.80-1.26
Hispanic	-	-	-	-
Parental crime	-	-	-	-

Number of prior arrests	1.17*** (40.49) C.I. 1.12-1.23	1.17*** (40.53) C.I. 1.12-1.23	1.17*** (40.53) C.I. 1.12-1.23	1.17*** (39.57) C.I. 1.12-1.23
Prior institutional living arrangement	-	-	-	-
Out of home placement ever	-	-	-	-
Instant Offense Type				
Sex offense		0.55** (7.77) C.I. 0.36-0.84	0.55** (7.77) C.I. 0.36-0.84	0.55** (7.60) C.I. 0.36-0.84
Drug offense		0.64*** (16.14) C.I. 0.51-0.79	0.64*** (16.14) C.I. 0.51-0.79	0.67*** (12.56) C.I. 0.54-0.84
Person offense		-	-	-
Property offense		1.48*** (21.38) C.I. 1.25-1.75	1.48*** (21.38) C.I. 1.25-1.75	1.49*** (21.85) C.I. 1.26-1.76
Neighborhood Social Disorg.				
Area drug sale density (ln)			-	-
Area % female HH w/ children (ln)			-	-
Area % vacant housing (ln)			-	-
Area % high school graduate (ln)			-	-

Contagion Effects				
Area juvenile count (ln)				-
Area any recidivism rate (ln)				-
Area drug recidivism rate (ln)				-
Area person recidivism rate (ln)				-
Area property recidivism rate (ln)				3.30*** (74.26) C.I. 2.51-4.33
Constant	0.10*** (2209.86)	0.10*** (1121.42)	0.10*** (1121.42)	1.43 (1.33)
Nagelkerke R2	0.01	0.03	0.03	0.06
Area under ROC curve				0.65*** C.I. 0.63-0.67
A gray box indicates a variable that was excluded from that model run. A dash indicates a variable that was allowed to enter that model but was not included by the stepwise procedure. Cell values indicate odds ratios. Wald statistic shown in parentheses. "C.I." indicates confidence interval at 95% confidence. *p<0.05, **p<0.01, ***p<0.005				

Discussion

Our findings from the univariate analysis are consistent with those of previous research. Social disorganization, drug and alcohol availability, family disorganization, parental behavior, minority group status, and the spatial concentration of delinquent peers in the environment have all been linked to juvenile offending in other studies. There is some indication that the removal of parents from the home due to criminal charges and the removal of youths from the home for any reason (delinquency, abuse or neglect) increase the likelihood of further delinquent behavior. This fracturing of the family can be interpreted in terms of the cause of family members being removed or the impact of such removal. It may be that an absence of effective parenting produced the removal, or it may be that removal from the home weakens bonds among family members, thus diminishing the potential for effective parenting.

Our multivariate analysis of recidivism, using all offense types, produced findings that are logically linked to previous research. Prior behavior is a strong predictor of future behavior, so the impact of number of prior arrests on recidivism was expected. Similarly, while not a risk measure, we would expect that youths that had been in an institutional setting immediately prior to their commitment to a community-based program were viewed by the court as higher risk cases than those youths whose offenses resulted in a community-based placement. Thus the positive relationship between prior institutional living arrangement and any recidivism variables was expected.

The impact of prior offense type was also expected. Sex offenders are known to recidivate at low rates in terms of new sex offenses, but to recidivate in other ways at levels commensurate with non-sex offenders (Caldwell, 2002; Hunter, Gilbertson, Vedros, and Morton, 2004). Our findings, however, indicate that a sex offense suppresses the likelihood of a subsequent delinquent offense of any kind. Although sex offenders recidivate with a variety of offenses, their offense behavior is less likely to be associated with other factors known to predict delinquency. Instead, research on this population shows that they are a heterogenous population (Hunter, Gilbertson, Vedros, and Morton, 2004).

The pattern for the other three offense types presents us with some indication that type of offense matters. In this case, drug offending and property offending, but not person offending, are associated with re-offending in general. Interestingly, both drug and property offending imply material gain, and it may be that success in obtaining desired material results conditions repetition of these acts.

Our finding that delinquent youths tend to specialize in the types of offenses they commit is consistent with previous research (Blumstein, 1988; Piquero, Paternoster, Maxerolle, Brame, and Dean, 1999; Armstrong, 2008). For each offense type, some degree of specialization was found. This tendency, however, is far stronger for those who committed drug offenses. A prior drug offense more than doubles the probability of a drug re-offense. Moreover, drug offending decreases the likelihood that a recidivating youth will commit some other type of an offense. These findings suggest that the causal mechanisms underlying drug offending differ from those influencing other types of offending.

Looking at environmental effects, we find that neighborhood disorganization was not influential after accounting for other explanatory variables, but that high rates of juvenile recidivism in the neighborhood surrounding individual youths greatly increased the likelihood of recidivism in individual offenders. This finding suggests a spatial contagion effect that is consistent with the impact of delinquent peers as a factor mediating the impact of neighborhood structural factors and parental behavior (Cattarello, 2000; Chung and Steinberg, 2006). To further investigate the influence of the contagion variables, we compared maps of the local spatial clustering of probability of recidivism from models 2 and 4 for each outcome variable. For the sake of brevity, we focus this discussion on just two of the outcomes: drug offense and person offense recidivism, as we feel these results were the most interesting of all modeled offense types.

Figure 5.1 shows two maps of juvenile delinquents, where each point in the map represents the home location of an individual juvenile and the color of the point indicates the presence and sign of significant local spatial clustering of the probability of drug offense recidivism. These maps were created from the probabilities generated by the logistic regressions of drug offense recidivism shown in Table 5.9. The map on the left is derived from probabilities generated from model 2 (including only individual and instant offense explanatory variables), while the map on the right is derived from probabilities generated from model 4 (also incorporating neighborhood social disorganization and contagion variables). Using these probability data, the G_i^* statistic (Getis and Ord, 1992; Ord and Getis, 1995) was calculated for each juvenile. The G_i^* statistic is a measure of the degree to which the observations within a distance *d* of observation *i* have values distinctly similar to, or different from, the global mean. Consider the spatial weights matrix $\{w_{ij(d)}\}$ such that $w_{ij(d)} = 1$ if location *i* is within distance *d* of location *j*, and $w_{ij(d)} = 0$ if it is not. In this study, *d* = 1 km. Let $W_i^* = \sum_j w_{ij}(d)$ and $S_{1i}^* = \sum_j w_{ij}^2(d)$, and let $\bar{z}$ and s^2 denote the sample mean and variance, respectively. The G_i^* statistic may be calculated as

$$G_i^*(d) = \frac{\sum_j w_{ij}(d) z_j - W_i^* \bar{z}}{s\{[(nS_{1i}^*) - W_i^{*2}]/(n-1)\}^{1/2}}$$

Significant local clusters of high and low probability of recidivism are colored red and blue, respectively. Due to the sheer number of points relative to the scale of the map, it is difficult to visually identify individual juveniles in the maps. However, the purpose of these maps is not to indicate the residence locations of individual juveniles, but rather to visualize the broad spatial patterns of clustering in the likelihood of drug offense recidivism as generated from the different logistic regression models.

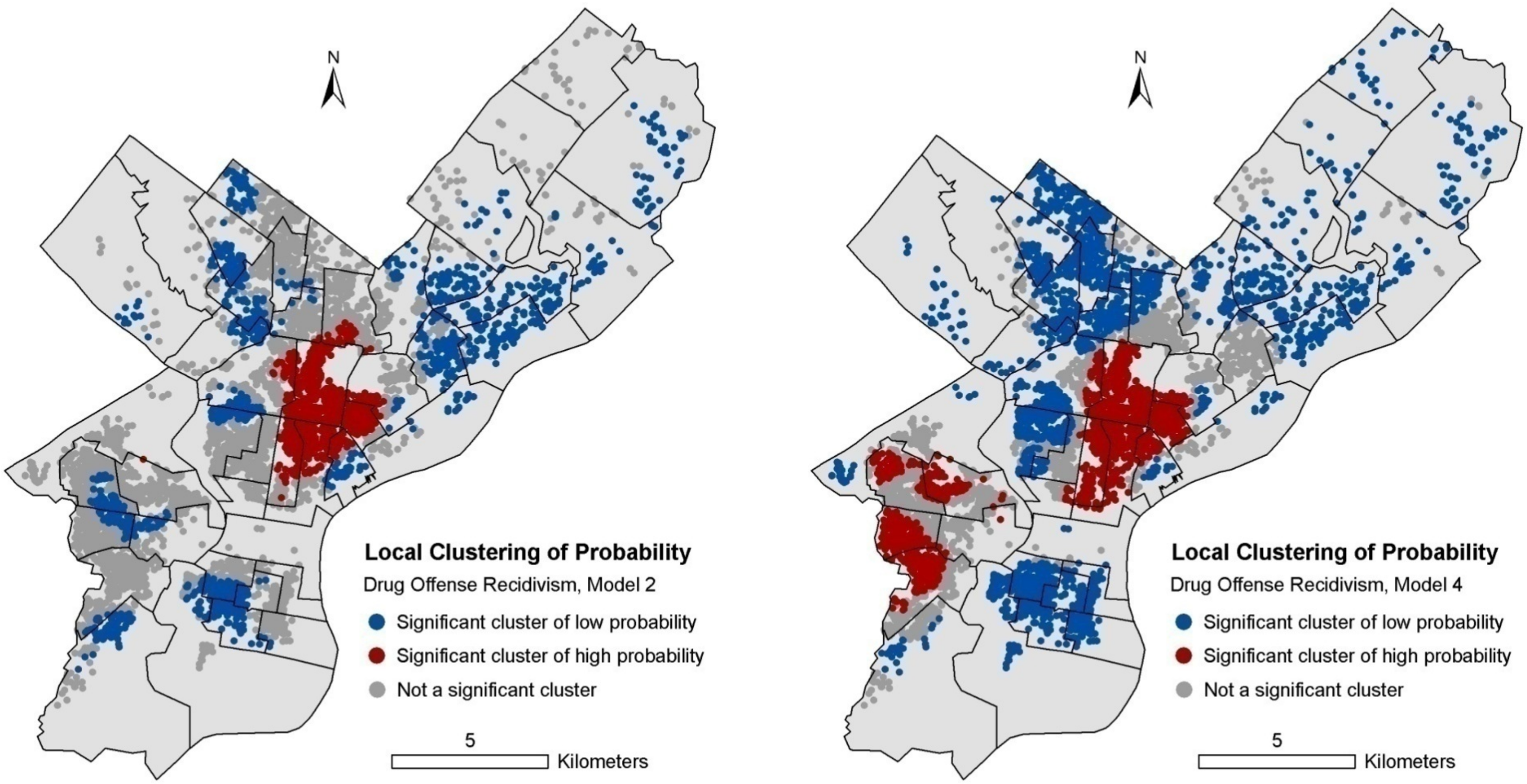
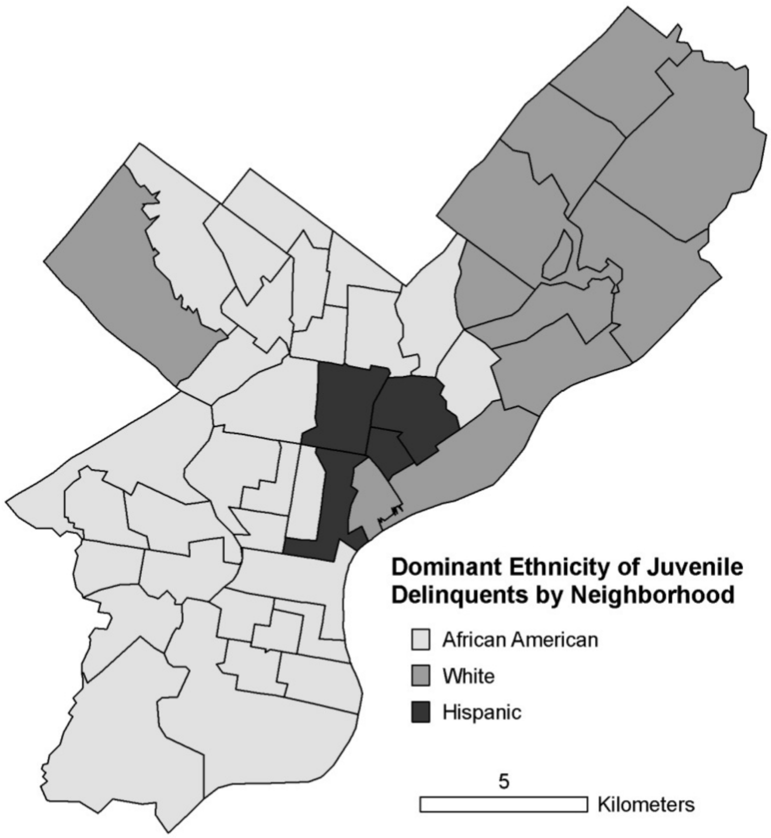

N
Local Clustering of Probability
Drug Offense Recidivism, Model 2
Significant cluster of low probability
Significant cluster of high probability
Not a significant cluster
5
Kilometers
N
Local Clustering of Probability
Drug Offense Recidivism, Model 4
Significant cluster of low probability
Significant cluster of high probability
Not a significant cluster
5
Kilometers

Figure 6.1

The map in Figure 6.1, left, suggests the substantial influence of ethnicity in the model. Compare this map with Figure 6.2, which shows the dominant ethnic group (African American, white, and Hispanic) of the juveniles residing in each neighborhood. Clearly, the cluster of high probability of drug offense recidivism observed in Figure 6.1, left, is spatially coincident with the concentration of Hispanic juveniles. This relationship is confirmed by Table 6.9, model 2, where being Hispanic increases the odds of drug offense recidivism by 1.65, while being white reduces the likelihood of drug offense recidivism by nearly half. The map in Figure 1, right, shows that the inclusion of the contagion variable, area drug recidivism rate, in the model allows for non-Hispanic juveniles to have a high estimated probability of drug offense recidivism. Thus, clusters of high probability of drug offense recidivism can be observed for several African-American neighborhoods in West Philadelphia. This is reflected in the odds ratios for Table 6.9, model 4, where Hispanic is not significant and the influence of being white is reduced.

Figure 6.2


Dominant Ethnicity of Juvenile Delinquents by Neighborhood
African American
White
Hispanic
5
Kilometers

A similar effect can be observed for models of person offense recidivism in Figure 6.3, which shows maps analogous to those in Figure 6.1, but for person offenses. In the map on the left, which shows local clusters of probabilities for model 2 in Table 6.10, the Hispanic neighborhood has a local cluster of low probability of recidivism, while many African American neighborhoods exhibit local clustering of high probability. In the map on the right, which shows local clusters of probabilities for model 4, African American neighborhoods exhibit local clustering of both high and low probability of recidivism, as do white neighborhoods.

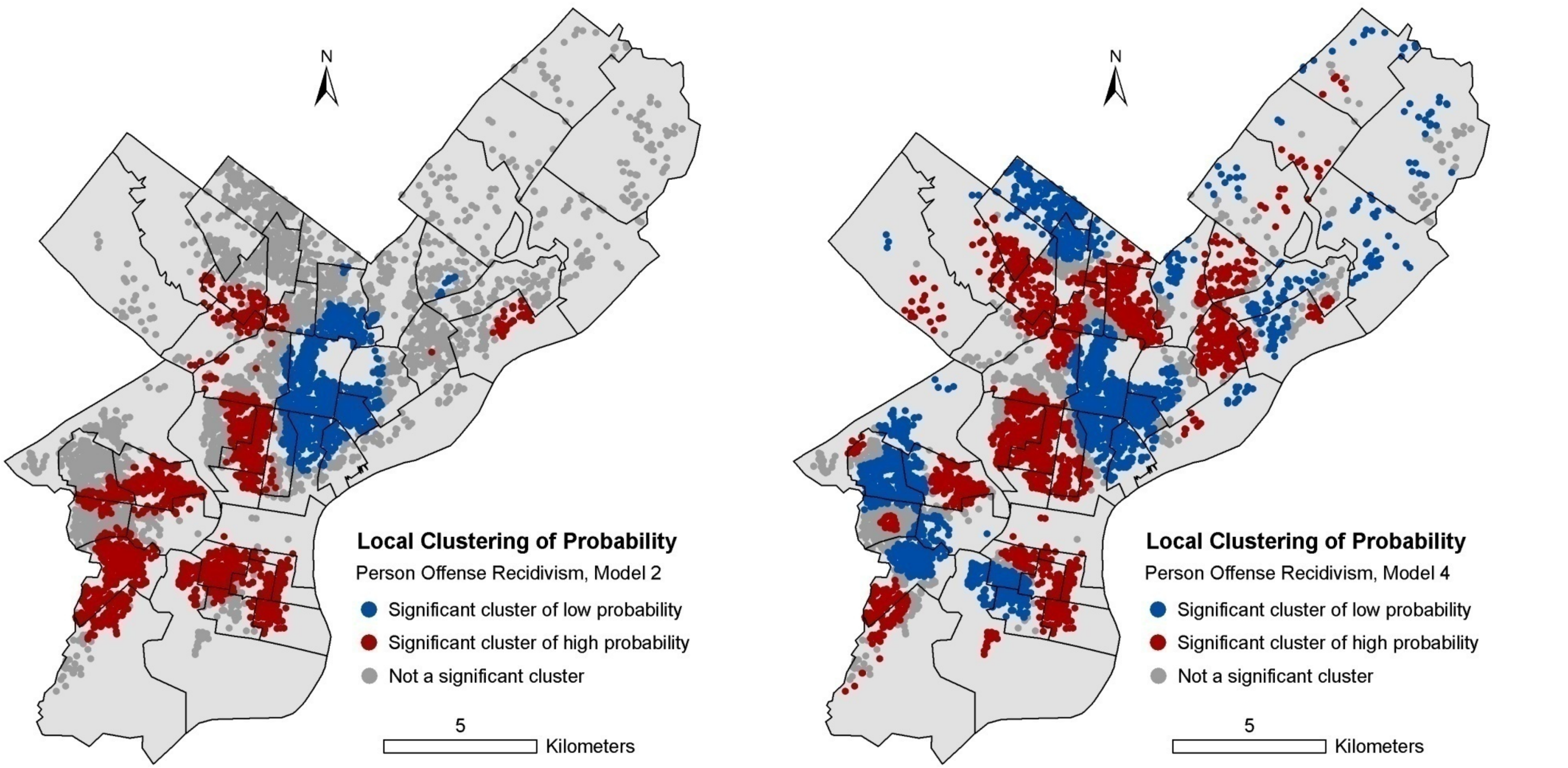

N
Local Clustering of Probability
Person Offense Recidivism, Model 2
Significant cluster of low probability
Significant cluster of high probability
Not a significant cluster
5
Kilometers
N
Local Clustering of Probability
Person Offense Recidivism, Model 4
Significant cluster of low probability
Significant cluster of high probability
Not a significant cluster
5
Kilometers

Figure 6.3

Further evidence of the important role of the contagion variables is illustrated by mapping the local spatial clustering of model residuals. Consider the maps shown in Figure 5.4, which show the G_i^* statistic calculated for the standardized residuals of the logistic regressions of drug offense recidivism. The residual for a given juvenile was calculated as $1 - P$ if the juvenile recidivated and $0 - P$ if the juvenile did not recidivate, where P is the probability of recidivism generated by the logistic regression. Residuals were converted to *z*-scores prior to calculation of the G_i^* statistic. The map on the left presents local clustering of the residuals derived from the model 2 in Table 6.9, while the map on the right is derived from model 4 in Table 6.9.

Figure 6.4, left, clearly shows local clusters of false positives (where the model incorrectly predicts drug offense recidivism for a juvenile who does not reoffend with a drug offense) and false negatives (where the model incorrectly predicts no drug recidivism). The most prominent cluster of false positives occurs in the Hispanic neighborhood, because the model is relying substantially on whether a juvenile is Hispanic to determine the likelihood of recidivism. There are also a number of scattered clusters of false negatives in African American neighborhoods, where the model incorrectly tends to predict an absence of drug offense recidivism. Figure 6.4, right, shows that once the contagion variable is entered into the model (Table 6.9, model 4), the significant local spatial clusters of false positives and false negatives all but disappear.

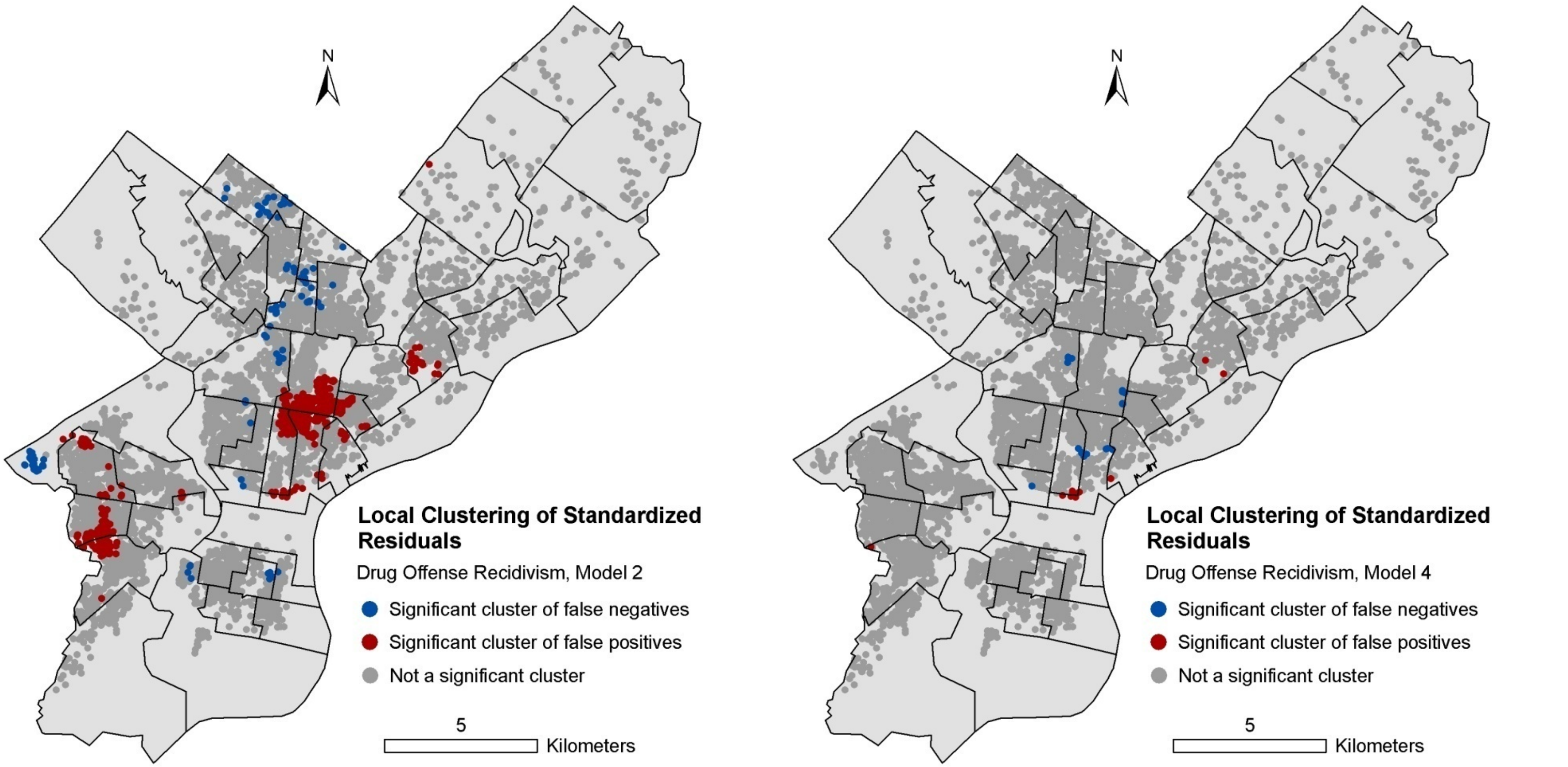
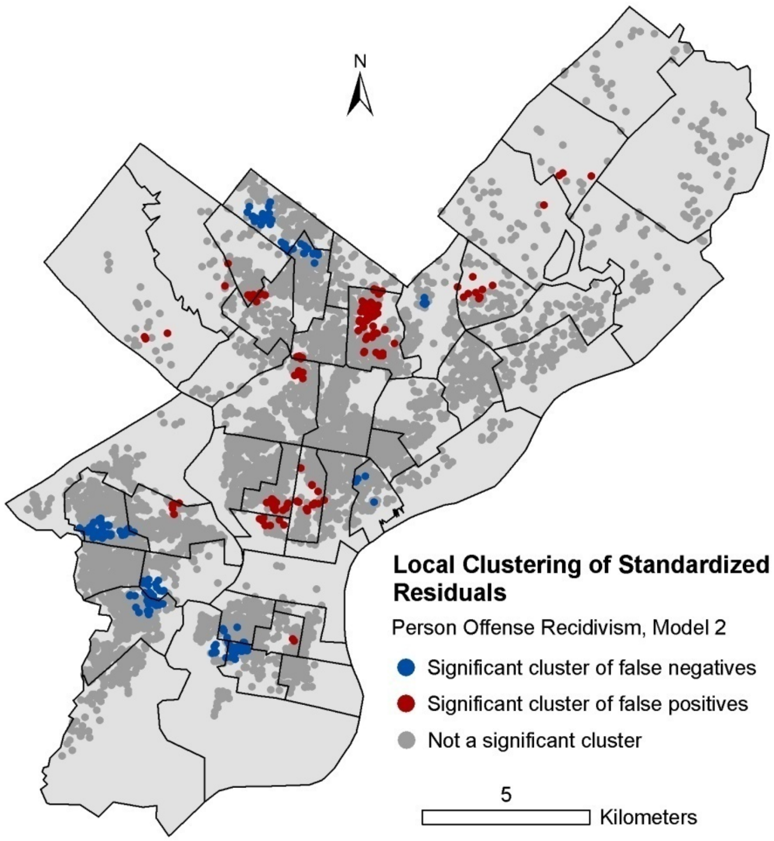
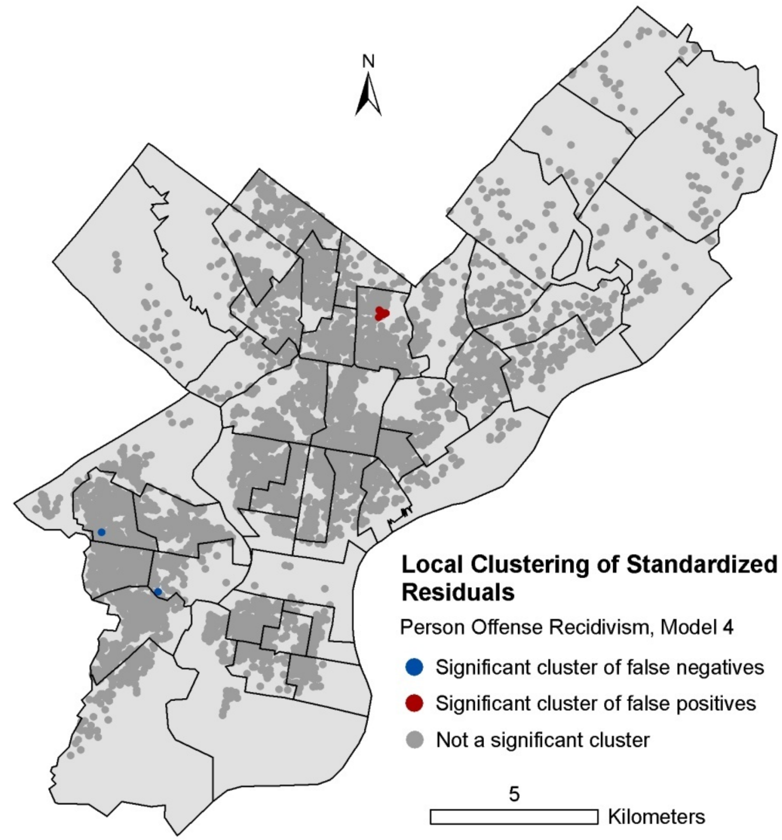

N
Local Clustering of Standardized Residuals
Drug Offense Recidivism, Model 2
Significant cluster of false negatives
Significant cluster of false positives
Not a significant cluster
5
Kilometers
N
Local Clustering of Standardized Residuals
Drug Offense Recidivism, Model 4
Significant cluster of false negatives
Significant cluster of false positives
Not a significant cluster
5
Kilometers

Figure 6.4

A similar pattern is evident for the analogous maps of the residuals of the models of person offense recidivism, shown in Figure 6.5. Figure 6.5, left, shows various clusters of false positives and false negatives, though in this case clusters of both types of errors occur primarily in African American neighborhoods. Figure 6.5, right, shows that the vast majority of the spatial clustering of these errors is not present in model 4.





Insert Figure 6.5

It appears that we are seeing causal patterns that are both offense-dependent and spatially-dependent – our subjects tend to specialize in neighborhoods that also specialize. Delinquent peer contagion, then, appears to be offense specific, facilitating offense specialization.

The extent of specialization among drug offenders, for example, relative to other offender types, indicates a relatively organized neighborhood structure that supports involvement in this type of delinquency. Indeed, Armstrong's (2008) finding that drug offending is, on balance, driven more by experience than family disruption or age (which at least implies maturity) suggests that drug offending is more likely to be influenced by environmental forces.

It is also possible that the contagion variables serve as proxies for other spatially dependent causal mechanisms that are not present in our models. For example, there may be some neighborhood characteristic that is absent from our models, such as neighborhood cohesion or trust in one's neighbors, that enhances or suppresses the likelihood of certain types of recidivism. We also acknowledge that it is possible that neighborhoods simply differ from one another in the nature of juvenile delinquency and recidivism. Consider, for example, the role of ethnicity in drug offense recidivism. It may be that the reasons a juvenile reoffends with a drug offense in African American neighborhoods and white neighborhoods are different than for juveniles in Hispanic neighborhoods – that the motivations and means by which juveniles become involved in drug activity differs across ethnic contexts.

CHAPTER 7
CONCLUSIONS

This study began with two broad goals and five research questions. Our goals were:

1. To investigate the usefulness of geospatial analyses and data mining of social science data to improve knowledge building capacities in juvenile justice, and

2. To examine the simultaneous effects of individual, neighborhood and program attributes on recidivism among delinquent youths.

These questions were stimulated by awareness that juveniles as individuals, during the times they are participating in court-ordered, community-based programs, are nested in neighborhoods and nested in programs. We noted that previous examinations of juvenile recidivism have limited themselves to individual and program attributes, largely with the intent of evaluating the impact of programs or developing risk prediction tools. The impact of neighborhood forces on recidivism is rarely considered, particularly in the development of risk assessment tools. And yet we know that tests of social disorganization and anomie theories which rely heavily on environmental attributes have for years shown the value of attending to the spaces in which young people live. This anomaly was mentioned by Kubrin and Stewart (2006) who noted the widespread belief that attributes of individuals are sufficient to explain recidivism. We found early on, however, that the analytic tools used in the social sciences were limited in their capacities to manage nested data, and that although environmental variables often appear in the literature, spatial dependency is rarely considered. So in addition to hierarchical linear modeling and logistic regression, we set out to explore the lessons to be learned from tools such as neural network analysis, co-clustering, spatial analysis and mapping, and geographically weighted regression. In addition, we created variables measuring the spatial density of adult and juvenile offending in order to measure the influence of local concentrations of crime and delinquency on individual youths.

Our research questions were:

1. What benefits are gained from analyzing social science data with geospatial and data mining analytic tools?

2. How do individual youth and family characteristics, program characteristics, and neighborhood characteristics interact to produce specific program outcomes, such as recidivism and placement in a more secure facility?

3. Why is recidivism more common in some neighborhoods than others?

4. Why are certain types of reoffending (offense type) more common in some neighborhoods than others?

5. To what extent is program impact a function of (constrained and enhanced by) neighborhood forces?

This final chapter pulls together the results of our attempts to answer these questions. Because our approach to this project was exploratory, some of our analytic decisions were also exploratory. Consequently, some findings are well supported while others are merely suggestive, requiring future research to test new hypotheses.

For example, the Plaid analysis reported in Chapters 2 and 3, found one group of individuals characterized by living in a largely African American neighborhood, where the proportion of residents with a high school education was considerably higher than average. Although race is a strong predictor of recidivism in our data, the addition of this educational characteristic identified a specific geographical area dominated by African Americans that was associated with a low recidivism rate. Educational level, then, may be an important Census variable to consider when employing neighborhood characteristics in predictions of recidivism risk.

In this chapter we discuss the different contributions of the methods used to this study of juvenile recidivism, followed by a set of conclusions about recidivism drawn from our findings. We begin with a discussion of the data.

Data-Related Decisions

This project involved the integration of different types of data: court and human services data on individual delinquent youths, police records of crime events, and a variety of spatial attributes of Philadelphia. We discuss here the two most critical sets of decisions made.

Working with Administrative Data

Certainly one of the greatest challenges we faced was deciding how to address a data set comprised of multiple cases for each youth, large amounts of missing data, and, within each case, relatively short periods of follow-up. The ProDES information system was conceived as a method for continuous monitoring of program outcome data. Consequently, each ProDES case was defined in terms of a program experience. These program commitments could be as the result of a new offense, a technical violation, a decision by the court that the youth had spent enough time at a program, or a request by the program to remove the youth. Each time a youth moved to a new program, a new case was created, and the entire data collection process began anew. Over time, some of these youths accumulated as many as ten program experiences.

Moreover, because delinquency program administrators were asked to collect clinical and individual performance data, variation in cooperation with the ProDES system produced large variations in the quality of data. Some programs faithfully collected data both at admission to the program and immediately prior to program discharge, others collected data at admission, but not at discharge, and still others collected no data at all.

Since the ProDES system was designed to provide continuous performance feedback to programs, the court and the agency that contracted with private providers, the follow-up period was restricted to six months following discharge from each program event. Re-offending was measured from the point of a new program commitment, and included adult arrests. The reason for this short follow-up period was that performance reports on each of approximately 110 programs were produced annually, with the expectation that the program and the larger system could make adjustments to program design and utilization. Reports then showed changes in performance measures over time.

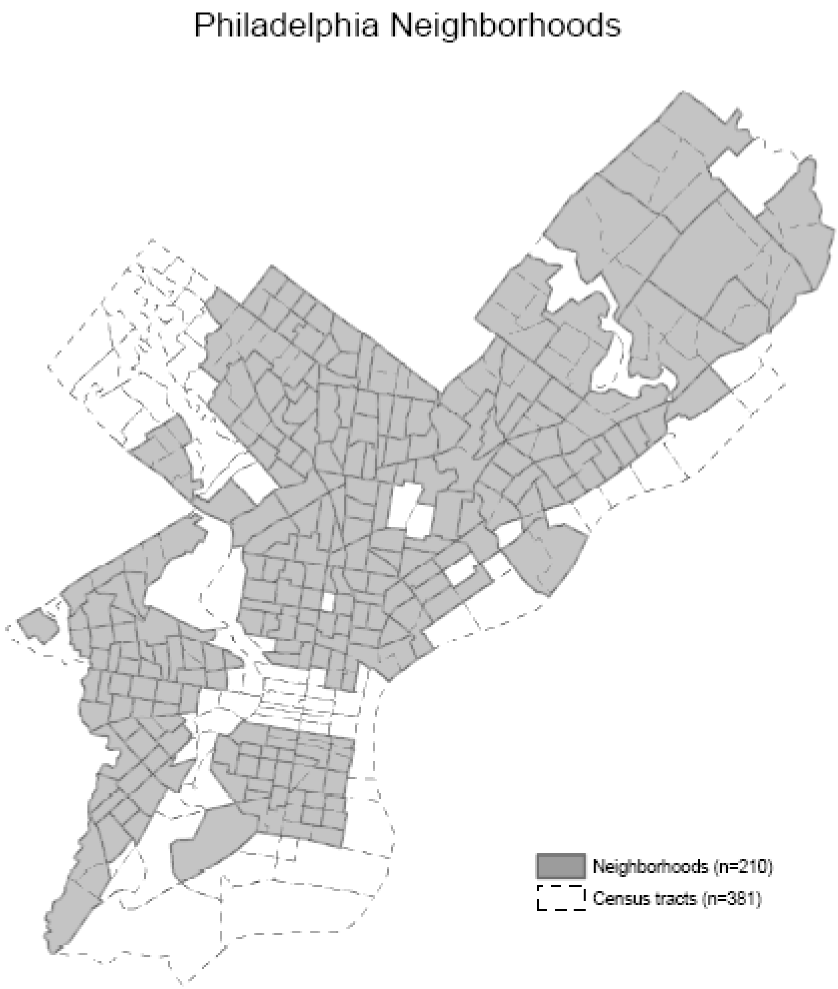
Because of the structure of these data around program commitment decisions, we decided to study the first time period in which the youth was committed to a community-based program. We chose to exclude time periods where the youth was confined in a residential facility outside the neighborhood of residence in order to study the impact of neighborhood on re-offending. Moreover, we chose to use only the first community-based program experience for a youth, since spatial analyses would be undermined by utilizing duplicates of the same youth.

Defining Neighborhood Boundaries

One critical step in our analysis was to create appropriate spatial units large enough to contain sufficient numbers of cases but small enough to represent reasonably a space familiar to each youth. As discussed in Chapter 4, we first utilized the neighborhood boundaries delineated by the Philadelphia Health Management Corporation (PHMC), which partitions the city into 45 neighborhood polygons, developed by means of focus groups. We found, however, that none of the neighborhood perception variables were significant predictors of youth recidivism. We believed that the level of aggregation was too large, resulting in the inclusion of several distinct populations within each neighborhood, thereby washing out important effects within the neighborhoods and reducing the variance between neighborhoods. We decided to use census tracts, then, for at least some of the analyses, arguing that census tracts would capture intra-neighborhood variance masked by the PHMC neighborhoods. The city of Philadelphia is divided into 381 Census tracts, which make for substantially smaller spatial units when compared to the 45 neighborhoods of the PHMC.

The next challenge was to take into account tracts that contained too few cases. An initial analysis of the number of cases per Census tract (n=318) indicated that the data was too sparse when nested in Census tracts. Of the 381 Census tracts, 83 contained zero cases, while an additional 114 tracts contained fewer than 20 cases each. Adjacent tracts were merged based on similarities in race and socioeconomic status, and taking into account major barriers such as rivers, railroads, and major highways, so that each spatial unit contained a minimum number of cases sufficient for supporting the reliable estimation of intra-neighborhood variance. The resultant neighborhood file contained 210 neighborhoods. Figure 7.1 (repeated here from Figure 4.1) illustrates the new neighborhood boundaries overlaid onto the 381 Census tracts of Philadelphia.

Figure 7.1. Philadelphia Neighborhood Delineations: Census Tracts vs. GIS Boundaries



Some merging of Census tracts was necessary in order to increase the number of cases. Thirty-eight neighborhoods had counts of juvenile addresses below twenty, with eleven juveniles being the fewest count per neighborhood. These neighborhoods were included in the study, and were not merged to increase the number of cases, in order to preserve theoretically interesting community attributes. We then aggregated Census and Police data to the 210 neighborhoods.

We believe that this method of maximizing spatial units and numbers of subjects within units is appropriate for the studies involving the impact of environmental attributes on individual behavior. For spatial analysis, however, other definitions of space, such as specifying fixed or variable distances from each subject, is more common, and was used in the spatial analysis reported in Chapter 5.

BENEFITS GAINED FROM ANALYZING SOCIAL SCIENCE DATA WITH GEOSPATIAL AND DATA MINING ANALYTIC TOOLS

The utilization of spatial analysis techniques to investigate the correlates of juvenile recidivism is rare in juvenile justice research. Recent research, however, examining the effects of community processes suggests that spacial factors exert a strong influence on the likelihood that juvenile offenders will reoffend. The inclusion of spatial indicators in models estimating juvenile recidivism permitted us to not only test the effects of community-level processes on recidivism, but also to investigate how individual-level effects can differ when controlling for neighborhood context. Further, by disaggregating juvenile recidivism by offense type, our research was able to understand the relationship between space and recidivism even more precisely.

Geographic analysis allows one to investigate the effects of environmental factors on behavior. There is simply no other way to address these environmental impacts without taking into account the characteristics of the neighborhood within which a person lives and spends time through geographic analysis. Our linear modeling studies reported in Chapter 4 reinforced the effects of environmental factors, and further aided in the development of our thinking about their differential effects on different types of offending.

Geographic analysis, however, allows for the investigation of spatial effects that do not take the form of direct impacts of geographic feature on a person. For example, our results in Chapter 6 suggest that social interaction among juveniles operating through spatial proximity drives criminal behavior. Without explicitly taking into account the locations of juveniles, there would be no way to discover such a pattern. One of our other major findings concerns the presence of nonstationarity, or heterogeneity, in our data. That is, subgroups of juveniles may be expected to exhibit different causal mechanisms for recidivist behavior. The identification of such subgroups may not be well understood from theory, but may be 'discovered' by examining patterns embedded in the data. Hence, subgroup analysis and pattern detection through data mining can lead to far more accurate models than conventional statistical modeling methods that ignore such heterogeneity.

The use of different analytic methods to study juvenile recidivism has provided an opportunity to both discover patterns in the data that we would otherwise have missed, and to test the reliability of findings. The sequence of chapters reflects the way in which research questions developed, with Chapter 6 examining the most interesting questions revealed by the other analyses.

The first contribution to this study came from sensitivity analyses, using neural network and association rule analyses. These preliminary analyses demonstrated the value of including spatial data in predictions of recidivism. In fact, it was this set of analyses that first indicated the impactof recidivism rates of other youths in close proximity to a juvenile. Ganypet33 (recidivism ratio of other offenders within 33 meters) and Probation (youth was on probation at the time of the instant offense) were found to be the variables that most increase the likelihood of recidivism.

A second preliminary study involved hot spot analysis. Here we employed a tool that provides mainly visual representations of patterns in the data. This hot spot analysis first revealed that recidivism, measured in terms of percent of delinquent youths that recidivated, was spatially dispersed. By conducting a local cluster analysis to classify cases into sub-regions of similar recidivism ratio, the groundwork was laid for observing characteristics of spatial dependency.

Co-clustering, using a method called Plaid, was employed to cluster our cases. What co-clustering avoids is the assumption that all cases can be clustered on the same set of attributes. Co-clustering, in fact, clusters simultaneously cases and attributes, producing what are called "layers'" and permits cases to belong to more than one layer. It was the preliminary findings of spatial dependency of different types of recidivism offenses, and the related differences among neighborhoods of the city that suggested to us the need to explore the possibility that different forces were operating on youths in different geographically defined areas of the city.

Neural networks analysis was again applied to testing the predictive value of co-clustering (Chapter 2). For recidivism in general, partitioning the data into layers of cases and variables did not produce more accurate or more sensitive models. For drug offense recidivism, however, the neural network models showed improved accuracy with the partitioned data, indicating that there are types of drug offense cases, and that drug offense recidivism is better predicted with different sets of predictors than by treating all drug offenders as a single group. The same was found for property offending. For person offense recidivism, however, we find that partitioning the data failed to improve predictive accuracy over a model representing all cases. It is for drug offending, however, that the benefits of co-clustering are greatest. This finding supports the conclusion that we emphasize in other chapters regarding the uniqueness of drug selling, relative to other types of recidivism offenses.

Our spatial analysis of the co-clusters, or layers (Chapter 3), further supported the hypothesis that different processes were operating in different neighborhoods, producing different patterns of re-offending. First, most of the layers were clearly associated with different neighborhoods, with one exception. One layer characterized by the youth's alcohol abuse and parental substance abuse was not spatially dependent. Layer 8, which has a very high recidivism rate, is also the area we associate most with drug selling. In Philadelphia, this is a predominantly Hispanic area, a fact we note in our other analyses. Second, this analysis provided evidence that risk factors such as alcohol and drug abuse, including parental drug abuse, and prior involvement with the justice system increase the likelihood that layer properties are associated with high recidivism.

It should not be surprising that we have concluded that space affects juvenile recidivism: criminologists have come to understand that place matters for all manner of criminal justice-related outcomes. This research serves to support the view that communities play an integral role in shaping choices made by juveniles and should be considered when estimating offense risk and when crafting programs of services to be delivered to different types of youths in different types of neighborhoods.

Predicting Recidivism

The following three questions are best discussed as a group. We did not address “placement in a more secure facility” in our analysis largely because our initial hot spot analyses showed youths placed in residential facilities due to technical violations to be concentrated in neighborhoods not characterized by high concentrations of re-offending. This spatial pattern suggested that decisions based on technical violations should be studied as a separate phenomenon. Moreover, our early findings of the spatial dependency of re-offense types created an opportunity to explore re-offending in ways not anticipated.

We draw five broad conclusions from the findings reported in Chapters 2 through 6:

1. Delinquent reoffending is spatially dependent rather than spatially diverse. This finding is strongest for drug offending,

2. Recidivism among delinquent adolescents spreads through a process of contagion and is offense-type specific (e.g. high concentrations of violent re-offending increases the chance that a delinquent youth will recidivate with a violent offense),

3. For some types of offending, especially drug selling, juveniles are likely to specialize. This specialization is likely to be influenced by opportunities, constraints and pressures present in the youth’s neighborhood,

4. Different spatial units and their unique problems require different causal models; social disorganization is not always a useful explanation of recidivism,

5. Drug selling is more easily predicted than other types of offending, and is more logically related to opportunities presented in the neighborhood, and

Clearly these conclusions are not discrete: they overlap in terms of spatial dependency, drug selling, and specialization.

From the spatial analysis reported in Chapter 6, we find that delinquent youths tend to specialize in committing offenses of a particular type, but that specialization is far more likely among drug offenders than youths committing non-drug offenses. Moreover, we contend that specialization is influenced by peer contagion. That is, youths tend to specialize in offenses in which other juveniles in their neighborhood specialize. This finding of spatially-dependent specialization suggests that there are neighborhood dynamics at play that we do not fully understand. The association between ethnicity and drug offending is particularly strong, and we note the impact of historical patterns of segregated Latino communities on drug selling discussed elsewhere (see, e.g. Bourgois, 2003). This finding is consistent with the argument raised by Baumer and Gustafson (2007) regarding instrumental crime. They found that evidence of "the crime generating effect of a high level of commitment to monetary success goals combined with a low level of commitment to legitimate means for pursuing such goals" (p. 651). Our finding with reference to Latino neighborhoods adds to this perspective by highlighting the potential for cultural responses to economic deprivation.

Other studies have found that peer influence plays a critical proximal role in decisions by youths to sell drugs. Their perceptions of the acceptability and profitability of drug dealing are influenced most directly by peers and young adults within their communities (Li et al. 1996; Ricardo 1994; Whitehead, Peterson, and Kaljee 1994). The spatial concentrations of drug selling are particularly strong, suggesting that youths in those areas are under significant pressure to participate in a business common to adult and juvenile neighbors. Their perceptions that "everyone is doing it" may be quite accurate in some of the neighborhoods we identify.

The predictors of person offense specialization, on the other hand, although also spatially-dependent, are less clear. Contrary to findings by Jacobs (2006), we did not find that the rate of lone-parent families was associated with person offending. Family disruption (referring to Sampson and Groves, 1989), in the form of parental criminality, did, however, predict person re-offending, but we do not know if family disruption is more likely in neighborhoods with high levels of person offending. On the other hand, parental criminality suggests an environment in which antisocial behavior can be learned. Aggression is known to be learned behavior (Bandura, 1986), suggesting support for a social learning explanation for the pattern we see. We do not know the nature of the parent's criminality, however.

The studies reported in Chapter 4, using HLM, provide evidence that neighborhood context, in the form of economic disadvantage and crime, is a significant predictor of juvenile recidivism when offense type is ignored and when examining only drug offense recidivism. Based on the measures of social disorganization used in the spatial analysis (Chapter 6), however, it appears that social disorganization does not play a significant role in juvenile recidivism, once the individual characteristics of juveniles are accounted for. It should be noted of course, that there are strong relationships among indicators of neighborhood social disorganization, such as crime and socioeconomic disadvantage, with indicators we captured at the individual level, such as race and public assistance. This is certainly the case in Philadelphia, where historical patterns of industrial development, residential settlement, and suburbanization have created a deeply segregated residential pattern with concentrated poverty in inner city, minority neighborhoods. We note that some previous studies that ascribe a causal effect to neighborhood social disorganization have used only spatially aggregated data (e.g. Sampson and Groves, 1989; Veysey and Messner, 1999), thus making it difficult to distinguish between the effect of those characteristics of social disorganization that may be measured at an individual level (e.g. race) versus those that are perhaps more characteristic of a neighborhood as a whole (e.g. vacant housing rate – while no individual lives in a vacant house, the rate of vacant housing in a neighborhood is indicative of its character).

Our findings are inconsistent with those of Little and Steinberg (2006) who conclude that "adolescents who sold the most drugs were more likely to live in contexts characterized by high physical and social disorder…" (Little and Steinberg, 2006: 378). Additionally, they found that drug activity increases violence within neighborhoods, net of their measures of social disorganization. Their conclusion that "traditional dimensions of social disorganization predict drug activity which, in turn, leads to higher levels of criminal violence", serves to tie drug and violent offending together in disadvantaged neighborhoods (Martinez et al., 2008: 866). We find, instead that areas of high concentrations of drug recidivism, where adult drug arrests are also concentrated, are not the same as those areas where violence is concentrated. This is not to say that drug crime and violence are spatially incompatible: instead we suggest that violence is more likely in locations where violence is normal and that these violent areas may be more disorganized than those where drug selling thrives.

It is likely that under conditions of specialization, different offense types require different causal explanations. If a single causal model was all that was needed, we would not expect to find spatial-dependency of offense types. Instead we find areas of the Philadelphia in which juvenile recidivists are exhibiting specialization of a particular offense type. This not only implies different causal models, but also suggests that neighborhood attributes must be part of the causal picture.

A number of our analyses produced finding that demonstrate that drug offending as a form of recidivism is different from person or property offending. In the first two-level HLM analysis, drug offenders, unlike youths committing person and drug re-offenses, were more likely to have committed a drug offense in the past, to be Hispanic, and to be on aftercare. Age at first arrest has varying effects across the HLM models in Chapter 4, with a positive correlation for drug recidivism, but a negative correlation with both person and property recidivism. These results can be interpreted to say that juveniles who offend earlier in their lives are more likely to recidivate with person and property offenses, while drug offenders come to the attention of the justice system at a later age. The cross-classified HLM analysis also showed that drug re-offenders are more likely than the other youths to reside in economically disadvantaged neighborhoods.

Other differences between recidivism involving drug offending and other offense types are that youths who commit drug offenses are much more likely to specialize in their offending, and to be influenced by high rates of drug offending by adults in the area immediately around them. Moreover, the logistic regression analysis reported in Chapter 6 finds that drug offending decreases the likelihood that a recidivating youth will commit some other type of an offense. These findings suggest that the causal mechanisms underlying drug offending differ from those influencing other types of offending.

The extent of specialization among drug offenders, relative to other offender types, indicates a relatively organized neighborhood structure that supports involvement in this type of delinquency. That is, opportunities to gain access to drugs must be present, and reinforcement of the behavior is likely. At the same time, the combination of poverty, Hispanic culture, high rates of adult drug selling and specialization, imply that other opportunities to engage in legitimate employment are less available than in other neighborhoods. These findings are consistent with the argument raised by Baumer and Gustafson (2007), in reference to drug selling, regarding "the crime generating effect of a high level of commitment to monetary success goals combined with a low level of commitment to legitimate means for pursuing such goals" (p. 651). Our finding with reference to Latino neighborhoods adds to this perspective by highlighting the potential for cultural responses to economic deprivation. Several studies conclude that the primary attraction of illicit drug selling is the potential income that is rarely attainable for youth in economically depressed neighborhoods (Reuter, MacCoun, and Murphy 1990).

Detecting program effects?

One general conclusion we have reached is that in the community, it appears that large programs that serve diverse populations of youths are less effective than small, more specialized programs that are tailored to the needs of youths being served and characteristics of the neighborhoods in which the youths reside. After controlling for individual-level characteristics, we found that as client capacity, or program size, increased, recidivism rates also increased. This finding held up even after we examined offense types separately. The question this analysis raised was whether or not programs added anything to the impact of sources of influence more common to the daily lives of our subjects.

To investigate this question, we employed HLM, using a two level-analysis that examined individual and program level attributes, and a cross-sectional analysis that added Program attributes to the initial Individual x Neighborhood analysis.

Our examination of Individual-Level x Program Level factors indicates that program-level correlates of juvenile recidivism have differing effects depending on recidivism offense type. Client capacity, or program size, increases levels of recidivism. But perhaps the most interesting finding is that program characteristics did not initially seem to have any effect on drug recidivists. While this may be initially surprising, it is important to remember that drug recidivists were the only youths in the neighborhood-level whose recidivism was markedly affected by neighborhood context.

Person offenders, in contrast, seem to be coming from families that are in disarray, in which parents are poor role models and may even be modeling the behavior that got these youths in trouble with the law. Such individual-level relationships with person recidivism are supportive of the relationships found between substance abuse services, group counseling, and staff-to-client ratio that were detected at the program-level. The finding that counseling and substance abuse services have positive effects on juveniles who are more likely to have criminal and drug-abusing parents is expected.

The correlation between staff-to-client ratio and person recidivism is at first counterintuitive, as a higher staff-to-client ratio (fewer staff members per client) is found to increase the likelihood of reoffending, but this relationship may be tied to the effects of group counseling. Group counseling is shown to reduce recidivism for person recidivists, and it may be that programs providing group counseling for its clients may naturally have higher staff-to-client ratios as a result of the nature of putting clients into group-based treatment.

Study Limitations

We note several limitations to our study. First, our measures of recidivism include a period of program participation and an additional six months. We have not included program effects in our models. We do know from other research being done on the same data set that program attributes are unrelated to drug re-offending (Harris, Grunwald, and Lockwood, unpublished). For person and property offending, on the other hand, there appears to be a program effect that should be included in future studies. Moreover, there is a large body of program evaluation research that supports the view that programs can reduce re-offending (Greenwood, 2008).

Recent studies of offense specialization have employed longitudinal analyses, examining several offense transitions over time. We have analyzed only one offense transition; thus we have not included prior offense transitions that may challenge our conclusion about offense specialization. Moreover, we have not examined changes in offending patterns with age or experience. Our analysis would also benefit from measures of parent-child relationships, as well as measures of neighborhood-family interactions. The influence of parents on recidivism, although mediated by peer influences (Chung and Steinberg, 2006), needs to be explored and expanded to include both positive (warmth, knowledge, and monitoring) and negative (criminality, substance abuse, physical abuse), as well as measures of family bonding (Giordano and Cernkovich, 1987),

Despite these limitations, however, our findings suggest lines of inquiry heretofore unexplored. Although other research has investigated deviant peer contagion, and still other research has examined offense specialization among delinquent youths, we have found that deviant peer contagion is spatially dependent, and that contagion is likely to be associated with specialization. These findings suggest that juveniles are drawn to specific types of offending by the spatially-bounded concentration of offense specialization among their peers. Research on causes of delinquency within neighborhoods, then, may produce more useful causal models than studies that ignore spatial concentrations of offense patterns.

REFERENCES

Ainsworth, J. W. (2002). Why does it take a village? The Mediation of neighborhood effects on educational achievement. *Social Forces, 81*(1), 117-152.

Akers, R. L. (1985). *Deviant behavior: A social learning approach* (3rd ed.). Belmont, CA: Wadsworth.

Akers, R.L. (1998). Social learning and social structure: A general theory of crime and deviance. Boston: Northeastern University Press.

Altschuler, D., & Brash, R. (2004). Adolescent and teenage offenders confronting the challenges and opportunities of reentry. *Youth Violence and Juvenile Justice, 2*, 72-87.

Anderson, A. L. (2002). Individual and contextual influences on delinquency: The role of the single-parent family. *Journal of Criminal Justice, 30*, 575-587.

Andrews, D. A., & Bonta, J. (1995). *Level of service inventory - Revised*. Toronto: Multi-Health Systems.

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis, 27*, 93-115.

Anselin, L. (1998). *Spatial econometrics: Methods and models*. Dordrecht: Kluwer Academic.

Anselin, L. (1999). Interactive techniques and exploratory spatial data analysis. In P. Longley, M. F. Goodchild, D. Maguire & D. Rhind (Eds.), *Geographical information systems: Principles, techniques, management and applications*. Cambridge: Geoinformational International.

Anthony, J. C. (2006). Deviant peer effects: Perspectives of an epidemiologist. In K. A. Dodge, T. J. Dishion & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth: Problems and solutions*. New York: The Guilford Press.

Armstrong, T. A. (2008). Are trends in specialization across arrests explained by changes in specialization occurring with age? *Justice Quarterly, 25*(1), 201-222.

Ashford, J., & LeCroy, C. W. (1990). Juvenile recidivism: A comparison of three prediction instruments. *Adolescence, 25*(98), 441-450.

Ashford, J. B., & Lecroy, C. W. (1988). Predicting recidivism: An evaluation of the Wisconsin juvenile probation and aftercare risk instrument. *Criminal Justice and Behavior, 15*(2), 141-151.

Baird, C. (1984). *Classification of juveniles in corrections: A model systems approach*. Madison, WI: National Council on Crime and Delinquency.

Bandura, A. (1969). *Principles of Behavior Modification*. Oxford: Holt, Rinehart & Winston.

Baumer, E. (2002). Neighborhood disadvantage and police notification by victims of violence. *Criminology, 40*(3), 579-616.

Baumer, E.P., & Gustafson, R. (2007). Social organization and instrumental crime: Assessing the empirical validity of classic and contemporary anomie theories. Criminology, 45, 617-664.

Baumer, E., Horney, J., Felson, R., & Lauritsen, J. L. (2003). Neighborhood disadvantage and the nature of violence. *Criminology, 41*(1), 39-72.

Bernard, T. J. (1982). *The cycle of juvenile justice*. New York: Oxford University Press.

Beyers, J. M., Loeber, R., Wikstrom, P. O., & Stouthamer-Loeber, M. (2001). What predicts adolescent violence in better-off neighborhoods? *Journal of Abnormal Child Psychology, 29*(5), 369-381.

Billy, J. O. G., & Moore, D. E. (1992). A multilevel analysis of marital and nonmarital fertility in the U.S. *Social Forces, 70*(4), 977-1011.

Blumstein, A. (1993). Making rationality relevant. *Criminology, 31*(1), 1-16.

Blumstein, A., Cohen, J., Das, S., & Moitra, S. (1998). Specialization and seriousness during adult criminal careers. *Journal of Quantitative Criminology, 4*, 303-345.

Boots, B. (2003). Developing local measures of spatial association for categorical data. *Journal of Geographical Systems, 5*(2), 139-160.

Bourgois, P. (2003). *In search of respect: Selling crack in El Barrio*. New York: Cambridge University Press.

Brooks-Gunn, J., Duncan, G. J., Klebanov, P. K., & Sealand, N. (1993). Do neighborhoods influence child and adolescent development? *The American Journal of Sociology, 99*(2), 353-395.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis, 28*, 281-298.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (2002). Geographically weighted summary statistics - A framework for localised exploratory data analysis. *Computers, Environment and Urban Systems, 26*(6), 501-524.

Burgess, R.L., & Akers, R.L. (1966). A differential association reinforcement theory of criminal behavior. *Social Problems*, 14, 128-147.

Bursik, R. J. (1980). The dynamics of specialization in juvenile offenses. *Social Forces, 58*, 851-864.

Bursik, R. J. (1988). Social disorganization and theories of crime and delinquency: Problems and prospects. *Criminology, 26*(4), 519-552.

Bursik, R. J., & Grasmick, H. G. (1993). *Neighborhoods and crime*. New York: Lexington Books.

Caldwell, M. F. (2002). What we do not know about juvenile sexual reoffense risk. *Child Maltreatment, 7*(4), 291-302.

Casetti, E. (1972). Generating models by the expansion method: Applications to geographic research. *Geographical Analysis, 4*, 81-91.

Casetti, E. (1997). The expansion method, mathematical modeling, and spatial econometrics. *International Regional Science Review, 20*, 9-32.

Catalano, R. F., Hawkins, J. D., Wells, E. A., Miller, J., & Brewer, D. (1991). Evaluation of the effectiveness of adolescent drug abuse treatment, assessment of risks for relapse, and promising approaches for relapse prevention. *Substance Use & Misuse, 25*(12), 1085-1140.

Cattarello, A. M. (2000). Community-level influences on individuals' social bonds, peer associations, and delinquency: A multilevel analysis. *Justice Quarterly, 17*(1), 33-60.

Centers, N. L., & Wiest, M. D. (1998). Inner city youth and drug dealing: A review of the problem. *Journal of Youth and Adolescence, 27*(3), 395-411.

Chaiken, M. R., & Johnson, B. D. (1988). *Characteristics of different types of drug involved youth*. Washington, DC: U.S. Department of Justice.

Chung, H. L., & Steinberg, L. (2006). Relations between neighborhood factors, parenting behaviors, peer deviance, and delinquency among serious juvenile offenders. *Developmental Psychology, 42*(2), 319-331.

Cohen, A. K. (1955). *Delinquent boys: The culture of the gang*. New York: Free Press of Glencoe.

Cohen, J. (1986). Research on criminal careers: Individual frequency rates and offence seriousness. In A. Blumstein, J. Cohen, A. Roth & C. Visher (Eds.), *Criminal careers and career criminals*. Washington, DC: National Academy Press.

Corrado, R. R., Vincent, G. M., Hart, S. D., & Cohen, I. M. (2004). Predictive validity of the Psychopathy Checklist: Youth Version for general and violent recidivism. *Behavioral Sciences & the Law, 22*(1), 5-22.

Cottle, C. C., Lee, R. J., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles. *Criminal Justice and Behavior, 28*(3), 367-394.

Coulton, C. J., Korbin, J. E., Su, M., & Chow, J. (1995). Community level factors and child maltreatment rates. *Child Development, 66*(5), 1262-1276.

Curry, D., & Spergel, I. (1998). Gangs homicide, delinquency, and community. *Criminology, 26*, 381-405.

Cybenko, G. V. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems, 2*(4), 303-314.

Daigle, L. E., Cullen, F. T., & Wright, J. P. (2007). Gender differences in the predictors of juvenile delinquency. *Youth Violence and Juvenile Justice, 5*(3), 254-286.

Darlington, R. B. (1990). *Regression and linear models*. New York: McGraw-Hill.

Dembo, R., Schmeidler, J., Nini-Gough, B., Sue, C. C., Borden, P., & Manning, D. (1998). Predictors of recidivism to a juvenile assessment center: A three year study. *Journal of Child and Adolescent Substance Abuse, 7*(3), 57-77.

Dembo, R., Wothke, W., Seeberger, W., Shemwell, M., Pacheco, K., & Rollie, M. (2002). Testing a longitudinal model of the relationships among hish risk youths' drug sales, drug use, and participation in index crimes. *Journal of Child & Adolescent Substance Abuse, 11*(3), 37-61.

Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistics Society, Series B, 39*(1), 1-38.

Dishion, T.J., Cipaldi, D., Spracklen, K.M., & Li, F. (1995). Peer ecology of male adolescent drug use. *Development and Psychopathology*, 7, 803-824.

Dishion, T. J., & Dodge, K. A. (2006). Deviant peer contagion in interventions and programs: An ecological framework for understanding influence mechanisms. In K. A. Dodge, T. J. Dishion & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth: Problems and solutions*. New York: The Guilford Press.

Dishion, T. J., Dodge, K. A., & Lansford, J. E. (2006). Findings and recommendations: A blueprint to minimize deviant peer influence in youth interventions and programs. In K. A. Dodge, T. J. Dishion & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth: Problems and solutions*. New York: The Guilford Press.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*(9), 755-764.

Dodge, K. A., Lansford, J. E., & Dishion, T. J. (2006). The problem of deviant peer influences in intervention programs. In *Deviant peer influences in programs for youth: Problems and solutions*. New York: The Guilford Press.

Duncan, G. J., & Raudenbush, S. W. (1999). Assessing the effects of context in studies of child and youth development. *Educational Psychologist, 34*, 29-41.

Duncan, R. D., Kennedy, W. A., & Patrick, C. J. (1995). Four-factor model of recidivism in male juvenile offenders. *Journal of Clinical Child Psychology, 24*(3), 250-257.

Duncan, T. E., Duncan, S. C., Okut, H., Strycker, L. A., & Hix-Sam, H. (2003). A multilevel contextual model of neighborhood collective efficacy. *American Journal of Community Psychology, 32*(3), 245-252.

Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review, 26*(3), 244-268.

Elliott, D. S., Huizinga, D. H., & Ageton, S. S. (1985). *Explaining delinquency and drug use*. Beverly Hills, California: Sage.

Elliott, D. S., Wilson, W. J., Huizinga, D. H., Sampson, R. J., Elliott, A., & Rankin, B. H. (1996). The effects of neighborhood disadvantage on adolescent development. *Journal of Research in Crime and Delinquency, 33*(4), 389-426.

Empey, L. T. (1978). *American delinquency, its meaning and construction*. Homewood, Ill.: Dorsey Press.

Fader, J. J., Harris, P. W., Jones, P. R., & Poulin, M. E. (2001). Factors involved in decisions on commitment to delinquency programs for first-time juvenile offenders. *Justice Quarterly, 18*(2), 323-341.

Fagan, J., & Freeman, R. B. (1999). Crime and work. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 25). Chicago: University of Chicago Press.

Farrington, D. P. (1986). Stepping stones to adult criminal careers. In J. Block & M. R. Yarrow (Eds.), *Development of antisocial and prosocial behavior* (pp. 359-384). New York: Academic Press.

Farrington, D. P., & Hawkins, J. D. (1991). Predicting participation, early onset and later persistence in officially recorded offending. *Criminal Behavior and Mental Health, 1*, 1-33.

Farrington, D. P., Snyder, H. N., & Finnegan, T. A. (1988). Specialization in juvenile court careers. *Criminology, 26*(3), 461-488.

Florax, R. J. G. M., & Van der Vlist, A. J. (2003). Spatial econometric data analysis: Moving beyond traditional models. *International Regional Science Review, 26*(3), 223-243.

Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *The Hare Psychopathy Checklist: Youth Version*. North Tonawanda, NY: Multi-Health Systems.

Fotheringham, A. S., & Brunsdon, C. (1999). Local forms of spatial analysis. *Geographical Analysis, 31*(4), 340-358.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (1998). Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A, 30*, 1905-1927.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (2000). *Quantitative geography: Perspectives on spatial data analysis*. London: Sage Publications.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (2002). Geographically weighted regression: The analysis of spatially varying relationships. Chichester: Wiley.

Frederick, B. (1999). *Factors contributing to recidivism among youth placed with the New York State Division For Youth.* Albany, NY: Office of Justice Systems Analysis, New York State Division of Criminal Justice Services.

Freisthler, B. (2004). A spatial analysis of social disorganization, alcohol access, and rates of child maltreatment in neighborhoods. *Children and Youth Services Review, 26*(803-819).

Freisthler, B., Gruenewald, P. J., Remer, L. G., Lery, B., & Needell, B. (2007). Exploring the spatial dynamics of alcohol outlets and child protective services referrals, substantiations, and foster care entries. *Child Maltreatment, 12*(2), 114-124.

Freisthler, B., & Merritt, D. H. (2006). Understanding the ecology of child maltreatment: A review of the literature and directions for future research. *Child Maltreatment, 11*(3), 263-280.

Freisthler, B., Needell, B., & Gruenewald, P. J. (2005). Is the physical availability of alcohol and illicit drugs related to neighborhood rates of child maltreatment? *Child Abuse & Neglect, 29*(9), 1049-1060.

Funk, S. J. (1999). Risk assessment for juveniles on probation. *Criminal Justice and Behavior, 26*(1), 44-68.
Furstenberg, F. F., Cook, T., D., Eccles, J., Elder, G. H., & Sameroff, A. (1999). *Managing to make it: Urban families and adolescent success*. Chicago: University of Chicago Press.

Garcia, R. Marie, Taylor, R. B., & Lawton, B. (2007). Impacts of violent crime and neighborhood structure on trusting your neighbors. *Justice Quarterly, 24(4),* 679-704.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*, 189-206.

Giordano, P. C., & Cernkovich, S. A. (1987). Family relationships and delinquency. *Criminology, 25*, 295-321.

Goodchild, M. F., Anselin, L., Applebaum, R. F., & Harthorn, B. H. (2000). Toward spatially integrated social science. *International Science Review, 23*(2), 139-159.

Gottfredson, S. D., & Moriarity, L. J. (2006). Statistical risk assessment: Old problems and new applications. *Crime & Delinquency, 52*(1), 178-200.

Gould, P. (1970). Is statistix inferens the geographical name for a wild goose? *Economic Geography, 46*, 439-448.

Greenwood, P. (2008). Prevention and intervention programs for juvenile offenders. *The Future of Children, 18*(2), 185-210.

Guerra, N. G., Kim, T. E., & Boxer, P. (2008). What works: Best practices with juvenile offenders. In R. D. Hoge, N. G. Guerra & P. Boxer (Eds.), *Treating the juvenile offender* (pp. 79-102). New York, NY: Guilford Press.

Harris, P. W., Grunwald, H., & Lockwood, B. (unpublished). Do programs matter? A system-wide examination of program correlates of juvenile recidivism. Department of Criminal Justice, Temple University.

Harris, P. W., & Jones, P. R. (1999). Differentiating delinquent youths for planning and evaluation. *Criminal Justice and Behavior, 26*(4), 403-434.

Harris, P. W., Welsh, W. N., & Butler, F. (2000). A Century of Juvenile Justice. In G. LaFree (Ed.), *Criminal justice 2000: The nature of crime: continuity and change* (Vol. 1, pp. 359-425). Washington, DC: National Institute of Justice.

Hawkins, R. P., Fremouw, W. J., & Reitz, A. L. (1982). A model useful in designing or describing evaluations of planned interventions in mental health. In A. J. McSweeny, W. J. Fremouw & R. P. Hawkins (Eds.), *Practical program evaluation in youth treatment*. Springfield, IL: Charles C. Thomas.

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic treatment of antisocial behavior in children and adults*. New York: Guilford Press.

Herrenkohl, T. I., Maquin, E., Hill, K. G., Hawkins, J. D., Abbott, R. D., & Catalano, R. F. (2000). Developmental risk factors for youth violence. *Journal of Adolescent Health, 26*(3), 176-186.

Hodges, K., & Kim, C.-S. (2000). Psychometric study of the Child and Adolescent Functional Assessment Scale: Prediction of contact with the law and poor school attendance. *Journal of Abnormal Child Psychology, 28*(3), 287-297.

Hoffman, J. P. (2002). A contextual analysis of differential association, social control, and strain theories of delinquency. *Social Forces, 81*(3), 753-785.

Hoge, R. D., & Andrews, D. A. (1996). *Assessing the youthful offender: Issues and techniques*. New York: Plenum Press.

Hoge, R. D., Andrews, D. A., & Leschied, A. W. (1996). An investigation of risk and protective factors in a sample of youthful offenders. *Journal of Child Psychology and Psychiatry, 37*(4), 419-424.

Howell, J. C. (1995). Guide for implementing the comprehensive strategy for serious, violent, and chronic juvenile offenders. Rockville, MD: Office of Juvenile Justice and Delinquency Prevention.

Huizinga, D., Loeber, R., Thornberry, T. P., & Cothern, L. (2000). *Co-occurrence of delinquency and other problem behaviors*. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention.

Humes, E. (1997). No matter how loud I shout: A year in the life of juvenile court. New York: Simon and Shuster.

Izzo, R. L., & Ross, R. R. (1990). Meta-analysis of rehabilitation programs for juvenile delinquents: A brief report. *Criminal Justice and Behavior, 17*, 134-142.

Jacob, J. (2006). Male and female youth crime in Canadian communities: Assessing the applicability of social disorganization theory. *Canadian Journal of Criminology and Criminal Justice, 48*(31-61).

Jones, K. (1991). Specifying and estimating multilevel models for geographical research. *Transactions of the Institute of British Geographers, 16*, 148-159.

Jones, P. R., Harris, P. W., Fader, J., & Grubstein, L. (2001). Identifying chronic juvenile offenders. *Justice Quarterly, 18*(3), 479-508.

Kahn, T. J., & Chambers, H. J. (1991). Assessing reoffense risk with juvenile sexual offenders. *Child Welfare, 70*(3), 333-345.

Katsiyannis, A., & Archwamety, T. (1997). Factors related to recidivism among delinquent youths in a state correctional facility. *Journal of Child and Family Studies, 6*(1), 43-55.

Klein, M. W. (1984). Offence specialisation and versatility among juveniles. *British Journal of Criminology, 24*, 185-194.

Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine Learning, 2*(3), 271-274.

Kowaleski-Jones, L. (2000). Staying out of trouble: Community resources and problem behavior among high-risk adolescents. *Journal of Marriage and the Family, 62*, 449-464.

Krysik, J., & Lecroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism among juvenile offenders. *Research on Social Work Practice, 12*(1), 71-81.

Kubrin, C. E., Squires, G. D., & Stewart, E. A. (2007). Neighborhoods, race, and recidivism: The community-reoffending nexus and its implications for African Americans. *Race Relations Abstracts, 32*(1), 7-37.

Kubrin, C. E., & Stewart, E. A. (2006). Predicting who reoffends: The neglected role of neighborhood context in recidivism studies. *Criminology, 44*(1), 165-197.

Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology, 1*, 451-476.

LaVigne, N., Mamalian, T., & Visher, C. (2003). *A portrait of prisoner reentry in Illinois*. Washington, DC: The Urban Institute.

Lazzeroni, L., & Owen, A. (2002). Plaid models for gene expression data. *Statistica Sinica, 12*(1), 61-86.

LeBaron, J. (2002). Examining the relative influence of community context on juvenile offender post-confinement recidivism. Unpublished Doctoral Dissertation, Rutgers University, Newark, NJ.

Lehmann, E. L., & Romano, J. P. (2005). *Testing Statistical Hypotheses* (Vol. 3). New York: Springer.

Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin, 126*, 309-337.

Liberman, A. (2007). Adolescents, Neighborhoods, and violence: Recent findings from the project on human development in Chicago neighborhoods. Washington, D.C.: National Institute of Justice.

Lipsey, M. W. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In J. McGuire (Ed.), *What works? Reducing reoffending*. New York: John Wiley and Sons.

Lipsey, M. W. (1999). Can intervention rehabilitate serious delinquents? *Annals of the American Academy of Political and Social Science, 564*(1), 142-166.

Little, M., & Steinberg, L. (2006). Psychosocial correlates of adolescent drug dealing in the inner city - Potential roles of opportunity, conventional commitments, and maturity. *Journal of Research in Crime and Delinquency, 43*(4), 357-386.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Loeber, R., Farrington, D. P., & Waschbusch, D. A. (1998). Serious and violent juvenile offenders. In R. Loeber & D. P. Farrington (Eds.), *Serious & violent juvenile offenders: Risk factors and successful interventions*. London: Sage Publications Inc.

Loeber, R., & LeBlanc, M. (1990). Toward a developmental criminology. In M. Tonry & N. Morris (Eds.), *Crime and Justice: A review of research* (Vol. 12). Chicago: University of Chicago Press.

Loeber, R., & Wikstrom, P.-O. H. (1993). Individual pathways to crime in different types of neighborhood. In D. P. Farrington, R. J. Sampson & P. O. Wikstrom (Eds.), *Integrating individual and ecological aspects of crime*. Stockholm, Sweden: Liber Forlag.

Ludwig, J., Duncan, G. J., & Hirschfield, P. (2001). Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment. *Quarterly Journal of Economics, 116*, 655-680.

Lussier, P., LeBlanc, M., & Proulx, J. (2005). The generality of criminal behavior: A confirmatory factor analysis of the criminal activity of sex offenders in adulthood. *Journal of Criminal Justice, 33*, 177-189.

Marczyk, G. R., Heilbrun, K., Lander, T., & DeMatteo, D. (2003). Predicting juvenile recidivism with the PCL:YV, MAYSI, and YLS/CMI. *International Journal of Forensic Mental Health, 2*(1), 7-18.

Markowitz, F. E., Bellair, P. E., Liska, A. E., & Liu, J. H. (2001). Extending social disorganization theory: Modeling the relationships between cohesion, disorder, and fear. *Criminology, 39*, 293-320.

Martinson, R. (1974). What works? Questions and answers about prison reform. *The Public Interest, 35*, 22-54.

Massey, D. S., & Denton, N. A. (1989). Hypersegregation in U.S. metropolitan areas: Black and Hispanic segregation along five dimensions. *Demography, 26*, 373-391.

Mazerolle, P. (1998). Gender, general strain, and delinquency: An empirical examination. *Justice Quarterly, 15*(1), 65-91.

McCord, J. (2003). Cures that harm: Unanticipated outcomes of crime prevention programs. *The Annals of the American Academy of Political and Social Science, 587*, 16-30.

McCord, J., Widom, C. S., & Crowell, N. A. (2001). Juvenile Justice, Juvenile Crime. Panel on Juvenile Crime: Prevention, Treatment, and Control. Washington, DC: National Academy Press.

Mears, D., & Aron, L. (2003). Addressing the needs of youth with disabilities in the juvenile justice system: The current state of knowledge. Washington, DC: National Council on Disability.

Mears, D., & Travis, J. (2004). *The dimensions, pathways, and consequences of youth reentry*. Washington, DC: The Urban Institute.

Mears, D. P., & Travis, J. (2004). Youth development and reentry. *Youth Violence and Juvenile Justice, 2*(1), 3-20.

Merton, R. K. (1995). Opportunity structure: The emergence, diffusion, and differentiation of a sociological concept, 1930s - 1950s. In F. Adler, W. S. Laufer & R. K. Merton (Eds.), *The Legacy of anomie theory*. New Brunswick, NJ: Transaction.

Miller, J., & Lin, J. (2007). Applying a generic juvenile risk assessment instrument to a local context. *Crime & Delinquency, 53*(4), 552-580.

Miner, M. H. (2002). Factors associated with recidivism in juveniles: An analysis of serious juvenile sex offenders. *Journal of Research in Crime and Delinquency, 39*(4), 421-436.

Minor, K. I., Hartmann, D. J., & Terry, S. (1997). Predictors of juvenile court actions and recidivism. *Crime & Delinquency, 43*(3), 328-344.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*(4), 674-701.

Moore, M. H. (1996). The future of the juvenile court: A theoretical framework that fits. *The Future of Children, 6*(3), 140-146.

Morenoff, J. D., Sampson, R. J., & Raudenbush, S. W. (2001). Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence. *Criminology, 39*(3), 517-559.

Myner, J., Santman, J., Cappelletty, G. G., & Perlmutter, B. F. (1998). Variables related to recidivism among juvenile offenders. *International Journal of Offender Therapy and Comparative Criminology, 42*(1), 65-80.

Nagin, D., & Paternoster, R. (1994). Personal capital and social control: The deterrence implications of a theory of individual differences in criminal offending. *Criminology, 32*, 581-606.

O'Campo, P., Xue, X., Wang, M.-C., & Caughy, M. O. B. (1997). Neighborhood risk factors for low birthweight in Baltimore: A multilevel analysis. *American Journal of Public Health, 87*(7), 1113-1118.

Oberwittler, D. (2004). A multilevel analysis of neighbourhood contextual effects on serious juvenile offending. *European Journal of Criminology, 1*(2), 201-235.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis, 27*, 286-306.

Osgood, D., & Chambers, J. M. (2004). Social disorganization outside the metropolis: An analysis of rural youth violence. *Criminology, 31*(8), 81-115.

Osgood, W., & Smith, G. (1995). Applying hierarchical linear modeling to extended longitudinal evaluations: The Boys Town follow-up study. *Evaluation Review, 19*, 3-28.

Parks, G. A., & Bard, D. E. (2006). Risk factors for adolescent sex offender recidivism: Evaluation of predictive factors and comparison of three groups based upon victim type. *Sexual Abuse: A Journal of Research and Treatment, 18*(4), 319-342.

Paternoster, R., Brame, R., Piquero, A., Mazerolle, P., & Dean, C. W. (1998). The forward specialization coefficient: Distributional properties and subgroup difference. *Journal of Quantitative Criminology, 14*, 133-154.

Pearson, K. (1904). On the theory of contingency and its relation to association and normal correlation. *Drapers' Company Research Memoirs, Biometric Series I*. London.

Piquero, A. R., & Brezina, T. (2001). Assessing Moffitt's adolescence-limited offending hypothesis. *Criminology, 39*(2), 353-370.

Piquero, A. R., Brezina, T., & Turner, M. G. (2005). Testing Moffitt's account of delinquency abstention. *Journal of Research in Crime and Delinquency, 42*(1), 27-54.

Piquero, A. R., Gibson, C. L., Tibbetts, S. G., Turner, M. G., & Katz, S., H. (2002). Maternal cigarette smoking during pregnancy and life-course-persistent offending. *International Journal of Offender Therapy and Comparative Criminology, 46*(2), 231-248.

Piquero, A. R., Paternoster, R., Mazerolle, P., Brame, R., & Dean, C. W. (1999). Onset age and offense specialization. *Journal of Research in Crime and Delinquency, 36*(3), 275-299.

Pullmann, M. D., Kerbs, J., Koroloff, N., Veach-White, E., Gaylor, R., & Sieler, D. (2006). Juvenile offenders with mental health needs: Reducing recidivism using wraparound. *Crime & Delinquency, 52*(3), 375-397.

Quinlan, R. (1993). C4.5: Programs for Machine Learning. San Mateo, CA: Morgan Kaufmann Publishers.

Quist, R. M., & Matshazi, D. G. M. (2000). The Child and Adolescent Functional Assessment Scale (CAFAS): A dynamic predictor of juvenile recidivism. *Adolescence, 35*(137), 181-192.

Raine, A., Brennan, P., & Mednick, S. A. (1994). Birth complications combined with early maternal rejection at age 1 year predispose to violent crime at age 18 years. *Archives of General Psychiatry, 51*, 984-988.

Rankin, B. H., & Quane, J. M. (2000). Neighborhood poverty and the social isolation of inner-city African American families. *Social Forces, 79*(1), 139-164.

Rankin, B. H., & Quane, J. M. (2002). Social contexts and urban adolescent outcomes: The interrelated effects of neighborhoods, families, and peers on African-American youth. *Social Problems, 49*(1), 79-100.

Rasmussen, L. A. (1999). Factors related to recidivism among juvenile sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 11*(1), 69-85.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd Ed.)*. Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., & Sampson, R. J. (1999). Ecometrics: Toward a science of assessing ecological settings, with applications to the systematic social observation of neighborhoods. *Sociological Methodology, 29*(1), 1-41.

Robertson, A. A. (2000). Final report: Comparison of community-based models for youth offenders part I: Program effectiveness and cost effectiveness. Mississippi State, MS: Social Science Research Center.

Roman, J., Kane, M., Turner, E., & Frazier, B. (2006). *Instituting lasting reforms for prisoner reentry in Philadelphia*. Washington, DC: The Urban Institute.

Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. New York: Wiley.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91*, 473-520.

Sampson, R. J. (1997). The embeddedness of child and adolescent development: a community-level perspective on urban violence. In J. McCord (Ed.), *Violence and childhood in the inner city*. Cambridge: Cambridge University Press.

Sampson, R. J. (2001). How do communities undergird or undermine human development? In A. Booth & N. Crouter (Eds.), *Does it take a village? Community effects on children, adolescents, and families*. Mahwah, NJ: L. Erlbaum.

Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory *The American Journal of Sociology, 94*(4), 774-802.

Sampson, R. J., & Laub, J. H. (1994). Urban poverty and the family context of delinquency: A new look at structure and process in a classic study. *Child Development, 65*, 523-540.

Sampson, R. J., Morenoff, J. D., & Earls, F. (1999). Beyond social capital: Spatial dynamics of collective efficacy for children. *American Sociological Review, 64*(5), 633-660.

Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing "neighborhood effects": Social processes and new directions in research. *Annual Review of Sociology, 28*, 443-478.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science, 277*(5328), 918-924.

Schreiber, K. (1992). The adolescent crack dealer: A failure in the development of empathy. *Journal of The AMerican Academy of Psychoanalysis, 20*(2), 241-249.

Schwalbe, C. S., Fraser, M. W., & Day, S. H. (2007). Predictive Validity of the joint risk matrix with juvenile offenders. *Criminal Justice and Behavior, 34*(3), 348-361.

Schwalbe, C. S., Fraser, M. W., Day, S. H., & Cooley, V. (2006). Classifying juvenile offenders according to risk of recidivism. *Criminal Justice and Behavior, 33*(3), 305-324.

Seltzer, M. H. (1994). Studying variation in program success: A multilevel modeling approach. *Evaluation Review, 18*(3), 342-361.

Shader, M. (2001). *Risk factors for delinquency: An overview*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Shaw, C. R., & McKay, H. D. (1942). Juvenile delinquency and urban areas : A study of rates of delinquency in relation to differential characteristics of local communities in American cities. In H. L. Voss & D. M. Peterson (Eds.), *Ecology, crime, and delinquency* (pp. xxxii, 451 p.). New York: Meredith.

Shaw, C. R., & McKay, H. D. (1969). Social factors in juvenile delinquency: A study of the community, the family, and the gang in relation to delinquent behavior. Washington, DC: National Commission on Law Observance and Enforcement.

Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J., Reuter, P., & Bushway, S. D. (1997). *Preventing crime: What works, what doesn't, what's promising. Report to the U.S. Congress*. Washington, D.C.: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.

Short, J.F., Jr., & Strodtbeck, F.L. (1965). *Group processes and gang delinquency*. Chicago: University of Chicago Press.

Sickmund, M. (2002). *Juvenile residential facility census, 2000: Selected findings*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Simcha-Fagan, O., & Schwartz, J. E. (1986). Neighborhood and delinquency: An assessment of contextual effects. *Criminology, 24*(4), 667-699.

Simmons, R. D. (2001). Ecological factors that predict recidivism of male juvenile offenders after release from institutional commitment. University of South Carolina.

Smith, W. R., & Monastersky, C. (1986). Assessing juvenile sex offenders risk for reoffending. *Criminal Justice and Behavior, 13*(2), 115-140.

Snyder, H. N. (2004). An empirical portrait of the youth reentry population. *Youth Violence and Juvenile Justice, 2*(1), 39-55.

Snyder, H. N., & Sickmund, M. (2006). *Juvenile offenders and victims: 2006 national report*. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention.

Solomon, A., Visher, C., LaVigne, N., & Osborne, J. (2006). Understanding challenges of prisoner reentry: Research findings from the Urban Institute's prisoner reentry portfolio. Washington, DC: The Urban Institute.

Spencer, M. B., & Jones-Walker, C. (2004). Interventions and services offered to former juvenile offenders reentering their communities: An analysis of program effectiveness. *Youth Violence and Juvenile Justice, 2*, 88-97.

Stanton, B., & Galbraith, J. (1994). Drug trafficking among African-American early adolescents - Prevalence, consequences, and associated behaviors and beliefs. *Pediatrics, 93*(6), 1039-1043.

Steinberg, L., Chung, H. L., & Little, M. (2004). Reentry of young offenders from the justice system. *Youth Violence and Juvenile Justice, 2*(1), 21-38.

Stoolmiller, M., & Blechman, E. A. (2005). Substance use is a robust predictor of adolescent recidivism. *Criminal Justice and Behavior, 32*(3), 302-328.

Struening, E. L., Wallace, R., & Moore, R. (1990). Housing conditions and the quality of children at birth. *Bulletin of the N Y Academy of Medicine, 66*(5), 463-478.

Sullivan, M. L. (2004). Youth perspectives on the experience of reentry. *Youth Violence and Juvenile Justice, 2*(1), 56-71.

Tanenhaus, D. S. (2004). *Juvenile justice in the making*. New York: Oxford University Press.

Tannenbaum, F. (1938). *Crime and the community*. New York: Columbia University Press.

Tibbetts, S. G., & Piquero, A. R. (1999). The influence of gender, low birth weight, and disadvantaged environment in predicting early onset of offending: A test of Moffitt's interactional hypothesis. *Criminology, 37*(4), 843-877.

Turner, M. G., Hartman, J. L., & Bishop, D. M. (2007). The effects of prenatal problems, family functioning, and neighborhood disadvantage in predicting life-course-persistent offending. *Criminal Justice and Behavior, 34*(10), 1241-1261.

Vazsonyi, A.T., Cleveland, H.H., & Wiebe, R.P. (2006). Does the effect of impulsivity on delinquency vary by level of neighborhood disadvantage? Criminal Justice and Behavior, 33, 511-541.

Veysey, B. M. (2008). Mental health, substance abuse, and trauma. In R. D. Hoge, N. G. Guerra & P. Boxer (Eds.), *Treating the juvenile offender*. New York, NY: Guilford Press.

Veysey, B. M., & Messner, S. F. (1999). Further testing of social disorganization theory: An elaboration of Sampson and Grove's community structure and crime. *Journal of Research in Crime and Delinquency, 36*, 156-174.

Vigdor, J. (2006). Peer effects in neighborhoods and housing. In K. A. Dodge, T. J. Dishion & J. E. Lansford (Eds.), *Deviant peer influences in programs for youth: Problems and solutions*. New York: The Guilford Press.

Visher, C., LaVigne, N., & Travis, J. (2004). Returning home: Understanding the challenges of prisoner reentry. Maryland pilot study: Findings from Baltimore. Washington, DC: The Urban Institute.

Webster, C., MacDonald, R., & Simpson, M. (2006). Predicting criminality? Risk factors, neighbourhood influence and desistance. *Youth Justice, 6*(1), 7-22.

White, J. L., Moffitt, T. E., Earls, F., Robins, L., & Silva, P. A. (1990). How early can we tell? Predictors of childhood conduct disorder and adolescent delinquency. *Criminology, 28*(4), 507-527.

Wiebush, R. G., Baird, C., Krisberg, B., & Onek, D. (1995). Risk assessment and classification for serious, violent, and chronic juvenile offenders. In J. C. Howell, B. Krisberg, J. D. Hawkins & J. J. Wilson (Eds.), *A Sourcebook: Serious, violent, and chronic juvenile offenders* (pp. 171-210). Beverly Hills, California: Sage.

Wiebush, R. G., McNulty, B., & Le, T. (2000). *Implementation of the Intensive Community-Based Aftercare Program*. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention.

Wikstrom, P.-O. H., & Loeber, R. (2000). Do disadvantaged neighborhoods cause well-adjusted children to become adolescent delinquents? A study of male juvenile serious offending, individual risk and protective factors, and neighborhood context. *Criminology, 38*(4), 1109-1142.

Wilkes, R., & Iceland, J. (2004). Hypersegregation in the twenty-first century. *Demography, 41*, 23-36.

Witten, I.H. and Frank, E. (2005). Data Mining: Practical machine learning tools and techniques", 2nd Edition, San Francisco: Morgan Kaufmann.

Wu, S., & Flach, P. A. (2002). *Feature selection with labelled and unlabelled data.* Paper presented at the ECML/PKDD workshop on Integration and Collaboration Aspects of Data Mining, University of Helsinki.

Yoshikawa, H. (1994). Prevention as cumulative protection: Effects of early family support and education on chronic delinquency and its risks. *Psychological Bulletin, 115*(1), 28-54.

APPENDICES

Appendix A: Variable Lists

Appendix B: Dynamic Clustering-Based Estimation of Missing Values in Mixed Type Data

Appendix C: Additional Maps

APPENDIX A: VARIABLE LISTS

PRODES

Case Information

Variable Number	Variable Name	Variable Label	
61	Qtr	Quarter at disposition	
69	month	Month of Dispostion	
70	year	Year of Disposition	
62	calyear	Calendar year	

Demographics – from record

Variable Number	Variable Name	Variable Label	
011	Sex		
012	Race		
072	Age		

Target Variables

Variable Number	Variable Name	Variable Label	
171	anyweap	First new offense involves a weapon	
172	anydrug	First new offense involves drugs	
173	anysex	First new offense is a sex offense	
174	ginprog	In-program Recidivism	
175	gpstprg6	Post-program Recidivism	
176	ganypet	Any Recidivism (In/Post-program)	
183	retpet	Has youth been placed because of a new petition?	
184	retvio	Has youth been placed because of a probation violation?	
185	retprg2	Has youth been committed to a residential program	
188	ginoff2	First In-program Offense	
189	gpstoff2	First Post-program Offense	
190	ganyoffX	First New Offense	

Spatial identifiers (these are not from ProDES but were added)

Variable Number	Variable Name	Variable Label	
004	tract_id	Tract ID	
005	neighb_id	Neighborhood ID	
182	nbh	Neighborhood Name	
191	fid_1		
192	fid_1_1		

Program – from record

Variable Number	Variable Name	Variable Label	
09	program	PROGRAM	
10	program2	Program (dual commit)	
64	altpclas	Alternate Program Classification	
65	altpcls2	Alternate Program Classification for dual commit	
66	progclas	Program Classification	
67	oohome	Program Setting	

Family – from records

Variable Number	Variable Name	Variable Label	
13	sibnum	NUMBER FULL NATURAL SIBS - FILE	
15	motdec	DECEASED MOTHER	
16	fatdec	DECEASED FATHER	
17	famhouse	FAMILY MEMBERS IN HOUSEHOLD	
18	livarr	JUVENILES LIVING ARRANGEMENT	
20	csrcinc2	PUBLIC ASSISTANCE INCOME	
21	hpabusem	HISTORY ALCOHOL ABUSE - MOTHER	
22	hpabusef	HISTORY ALCOHOL ABUSE - FATHER	
23	hpdabusm	HISTORY DRUG ABUSE - MOTHER	
24	hpdabusf	HISTORY DRUG ABUSE - FATHER	
25	hpcrimm	HISTORY CRIMINALITY - MOTHER	
26	hpcrimf	HISTORY CRIMINALITY - FATHER	
27	sibarr	ANY SIBLINGS ARRESTED	
28	referral	ANY RECORD REFERRAL DHS	
29	hfmviole	HISTORY FAMILY VIOLENCE - EVER	

Youth Social History - Record

Variable Number	Variable Name	Variable Label	
30	svvict	Victim of sexual violence?	
31	intvio	Victim of violence in instit/resid. setting?	
32	juvkids	NUMBER OF CHILDREN FOR JUVENILE	
33	jhisaa	ANY HISTORY OF ALCOHOL ABUSE	
34	jhisda	ANY HISTORY OF DRUG ABUSE	
35	drgarr	Number of juv. arrests for drug offense	
36	jhismh	HISTORY OF MENTAL HEALTH PROBLEMS	
52	edstdisp	EDUCATIONAL STATUS AT DISPOSITION	
54	juvsped	JUVENILE IN SPECIAL ED	
74	IQ	IQ Classification Standardized by Test	

Program Discharge

Variable Number	Variable Name	Variable Label	
56	nincreps	Number incident reports	
57	apclient	Appropriate client	
58	drisktot	Discharge -- total risk score	
59	dneedtot	Discharge -- TOTAL NEEDS SCORE	
60	retprog	Has the juv. been committ. to a more secure/inten. program since disch.?	
158	disyear	Year of Discharge -calendar	
159	fydisyr	Fiscal Year of Discharge	
160	postprog	Next Program	
161	matchok2	Case for transfer to discharge scales	
162	tinprog	Time in Program (weeks)	
163	gtinprog	Time in Program (Grouped)	
164	progcomp	Did Youth Complete Program	
165	progtime	Intake to Discharge (weeks)	

Offense History

Variable Number	Variable Name	Variable Label	
45	prjda	NUMBER PRIOR DELINQUENCY ARRESTS	
46	prjda12	NUMBER PRIOR DELINQUENCY ARRESTS 12 MONTHS	
		PRIOR CHARGES SERIOUS PERSONAL OFFENSES	
47	pcspo	NUMBER PRIOR PROBATIONS	
		# OF OUT OF HOME PLACEMENTS	
48	nopprob	TOTAL TIME SPENT IN PLACEMENTS	
49	oohpnum	NUMBER RUNAWAYS/ESCAPES	
50	oohptime	Age At First Arrest	
51	prenum	Recoded prior arrests	
78	age1arr	Recoded recent prior arrests	
80	gprjda	Recoded # prior probations	
81	gprjda12	Recoded prior # out-home placements	
82	gnopprob	Recoded # prior charges serious personal offenses	
83	goohpnum	Out of home time in months	
84	gpcspo		
85	gohptime		

Instant Offense

Variable Number	Variable Name	Variable Label	
37	violate	VIOLATION OF PROBATION OR COURT ORDER	
38	totcharg	TOTAL # OF CHARGES	
39	victinj	INJURY TO VICTIM	
40	weapinv	WEAPON INVOLVED	
41	sexoff	ANY CHARGES FOR SEXUAL OFFESE	
42	otharrt	Number of others arrested for the same offense	
43	suplevio	SUPERVISION LEVEL TIME OF INSTANT OFFENSE	
44	outcome	OUTCOME OF CASE	
75	gtcharge	Number of charges	
76	goffc1	Offense Type	

Risk and Needs Assessment – Program reported

Variable Number	Variable Name	Variable Label	
86	riskclas	CHAID Risk Classification	
87	rskmiss	Number of intake risk items missing	
88	riskgood	Number of intake risk items present	
89	np1miss	Number of intake needs items missing	
90	np1good	Number of intake needs items present	
91	np1tot	Total of Part 1 Intake Needs	
92	nedsub2	Calculated Intake Education Needs Total	
93	npeersub	Calculated Intake Peer Needs Total	
94	nhlthsub	Calculated Intake Health Needs Total	
95	nsexsub	Calculated Intake Sex Needs Total	
96	pt2sub	Total of Part 2 Intake Needs	
97	ntot	Needs Total Score	
98	need_2	Intake Needs Category (Mean Based)	

Clinical Self Report Scales

Variable Number	Variable Name	Variable Label	
099	sesteem2	Self Esteem (Mean Based)	
100	ses	Intake Self-Esteem Score (ProRated)	
101	esteem_2	Intake Self Esteem Category (Mean Based)	
102	values1	Values Score (0-38) Not Adjusted For Missing	
103	values2	Intake Values Score (Adjusted For Missing)	
104	rdvalue	Re-directioned Intake Values Score	
105	grdvalue	Grouped Intake Values	
106	sattach	Intake School Attachment Score	
107	atteach	Intake Attachment to Teachers	
108	schcomit	Intake School Commitment	
109	peropp	Intake Perceived Opportunities	
110	conarr	Intake Consequences of Arrest	
111	schinv	Intake School Involvement	
112	comminv	Intake Community Involvement	
113	gsattach	Intake School Attachment Score	
114	gatteach	Intake Attachment to Teachers	
115	gschomit	Intake School Commitment	
116	gperopp	Intake Perceived Opportunities	
117	gconarr	Intake Consequences of Arrest	
118	gschinv	Intake School Involvement	
119	gcomminv	Intake Community Involvement	
120	fcartr	Intake Caring and Trust Score	
121	fidsup	Intake Identity Support Score	
122	fconsup	Intake Control and Supervision Score	

123	fintcom	Intake Intimacy and Communication Score	
124	fincomm	Intake Instrumental Communication Score	
125	gfcartr	Intake Caring and Trust	
126	gfidsup	Intake Identity Support	
127	gfconsup	Intake Control and Supervision	
128	gfintcom	Intake Intimacy and Communication	
129	gfincomm	Intake Instrumental Communication	
130	Zses	Zscore: Intake Self-Esteem Score (ProRated)	
131	Zvalues2	Zscore: Intake Values Score (Adjusted For Missing)	
132	Zsattach	Zscore: Intake School Attachment Score	
133	Zatteach	Zscore: Intake Attachment to Teachers	
134	Zschcomit	Zscore: Intake School Commitment	
135	Zperopp	Zscore: Intake Perceived Opportunities	
136	Zconarr	Zscore: Intake Consequences of Arrest	
137	Zschinv	Zscore: Intake School Involvement	
138	Zcomminv	Zscore: Intake Community Involvement	
139	Zfcartr	Zscore: Intake Caring and Trust Score	
140	Zfidsup	Zscore: Intake Identity Support Score	
141	Zfconsup	Zscore: Intake Control and Supervision Score	
142	Zfintcom	Zscore: Intake Intimacy and Communication Score	
143	Zfincomm	Zscore: Intake Instrumental Communication Score	
144	ZSco01	Zscore(ses) Intake Self-Esteem Score (ProRated)	
145	ZSco02	Zscore(values2) Intake Values Score (Adjusted For Missing)	
		Zscore(sattach) Intake School Attachment Score	
146	ZSco03	Zscore(atteach) Intake Attachment to Teachers	
147	ZSco04	Zscore(schcomit) Intake School Commitment	
148	ZSco05	Zscore(peropp) Intake Perceived Opportunities	
149	ZSco06	Zscore(conarr) Intake Consequences of Arrest	
150	ZSco07	Zscore(schinv) Intake School Involvement	
151	ZSco08	Zscore(comminv) Intake Community Involvement	
152	ZSco09	Zscore(fcartr) Intake Caring and Trust Score	
153	ZSco10	Zscore(fidsup) Intake Identity Support Score	
154	ZSco11	Zscore(fconsup) Intake Control and Supervision Score	
155	ZSco12	Zscore(fintcom) Intake Intimacy and Communication Score	
156	ZSco13	Zscore(fincomm) Intake Instrumental Communication Score	
157	ZSco14		

Recidivism

Variable Number	Variable Name	Variable Label	
166	inprog	New In-program Petitions	
167	postprg6	New Post-program Petitions	
168	anypet	Total New Petitions (In/Post-program)	
169	inoff1	In-program first offense	
170	pstoff1	Post-program first offense	
171	anyweap	First new offense involves a weapon	
172	anydrug	First new offense involves drugs	
173	anysex	First new offense is a sex offense	
174	ginprog	In-program Recidivism	
175	gpstprg6	Post-program Recidivism	
176	ganypet	Any Recidivism (In/Post-program)	
177	gintime1	Time to In-program 1st Petition (weeks)	
178	gpstime1	Time to 1st Post-program Petition (weeks)	
179	gpet1tim	Time to 1st Petition (weeks)	

PROGRAM DESIGN INVENTORY (PDI)

Organizational/Classification Information

Address
Phone/Fax/Internet Address
Program Mission
Date Program Began
DHS Classification
RTF Certification
Average Length of Stay (Months)
City Population Size
Total # of Program Evaluations Per Year
Does Program Have At Least One Outcome Evaluation Each Year
Does Program Have At Least One Financial Evaluation Each Year
Does Program Have At Least One Policy Evaluation Each Year
Date Program Began with CJRI
CJRI Classification
Program Structure
Security Level of Program
Does Program Conduct Follow-Up
License Information

Target Population

Minimum Age of Target Population
Maximum Age of Target Population
Total Client Capacity of Program
Maximum # of Philadelphia Clients
Gender(s) of Target Population
Are Services Delivered Coed
Geographic Area Served
Philadelphia Probation District(s) Served
Problems/Needs of Target Population
Offense Behavior of Target Population
Behavior That Is Grounds for Program Rejection
Behavior That Is Grounds for Program Removal

Program Objectives

Objective Classification
Program Objective
Priority-Level
Time Frame

Program Services

Program Activity
Service Location(s)
Delivery Time
# of Hrs Per Week
Is a 3rd Part Vendor Used
Service Notes

Staff Background

of Direct-Care Female Staff
# of Direct-Care Male Staff
Total # of Direct-Care Staff
# of Clients Per Direct-Care Staff
Average Age of Direct Care Staff
# of Licensed Social Workers
# of Licensed Staff (excluding LSW)
# of Training Hrs w/i 90 Days of Employment
# of In-Service Training Hrs Per Year
# of White Direct-Care Staff
# of Black Direct-Care Staff
# of Hispanic Direct-Care Staff
# of Asian Direct-Care Staff
# of Direct-Care Staff w/ Other Racial Backgrounds
Staff Turnover During Previous Calendar Year
Does the Program Use Volunteers
Education Requirements
Is Related Work Experience Required
Foreign Language(s) Fluently Spoken by Direct-Care Staff
Language Spoken
# of Staff Who Speak the Language

Appendix B: Dynamic Clustering-Based Estimation of Missing Values in Mixed Type Data

Vadim Ayuyev[1], Joseph Jupin[2], Philip W. Harris[3] and Zoran Obradovic[2],

[1] FN1-KF Department, Bauman Moscow State Technical University (Kaluga Branch), Bazgenova Str. 2, Kaluga, 248600, Russian Federation
Vadim.Ayuyev@gmail.com
[2] Center for Information Science and Technology, Temple University, 303 Wachman Hall, 1805 N. Broad St., Philadelphia, PA 19122, USA
Joejupin@temple.edu, Zoran@ist.temple.edu
[3] Department of Criminal Justice, Temple University, 303 Wachman Hall, 1805 N. Broad St., Philadelphia, PA 19122, USA
Phil.Harris@temple.edu

Abstract. The appropriate choice of a missing data imputation method becomes especially important when the fraction of missing values is large and the data are of mixed type. The proposed dynamic clustering imputation (DCI) algorithm relies on similarity information from shared neighbors, where mixed type variables are considered together. When evaluated on a public social science dataset of 46,043 mixed type instances with up to 33% missing values, DCI resulted in more than 20% improved imputation accuracy over Multiple Imputation, Predictive Mean Matching, Linear and Multilevel Regression, and Mean Mode Replacement methods. Data imputed by 6 methods were used for test of NB-Tree, Random Subset Selection and Neural Network-based classification models. In our experiments classification accuracy obtained using DCI-preprocessed data was a lot better than when relying on alternative imputation methods for data preprocessing.

Keywords: Data pre-processing, data imputation, clustering, classification.

1 Introduction

A common approach to analyzing data with missing values is to remove attributes and/or instances with a large fraction of missing values. Such data preprocessing is appealing because it is simple and also reduces dimensionality. However, this is not applicable when missing values cover a lot of instances, or their presence in essential attributes is large [1].

Another common and practical way to address the problem of missing values in data is to replace them as estimates derived from the non-missing values by a linear function. The missing attribute *j* from an instance *i*, denoted as $x_{i,j}^{ms}$, is estimated as:

$$x_{i,j}^{ms} = f\left(x_{1,j}, x_{2,j}, \ldots, x_{p,j}, \ldots, x_{P_j,j}\right) , \qquad (1)$$

where f is a linear function of P_j variables; P_j is the number of instances in the data with non-missing values for attribute j; and $x_{p,j}$ is a non-missing attribute j from an instance p.

A special case of (1), which is simple, fast, and often provides satisfactory results when the number of missing values is relatively small and their distribution is random, is mean (or mode for categorical attributes) value based imputation:

$$x_{i,j}^{ms} = \frac{1}{P_j}\sum_{p=1}^{P_j} x_{p,j} \quad . \qquad (2)$$

The limitation of mean value based imputation and its variations is its focus on a specific variable without taking into account the overall similarities between instances. For example, consider the following 5 data points with 6 attributes, where a categorical attribute (fifth column) is missing one value (denoted as "*ms*"):

$$\begin{bmatrix} 1 & 10.2 & 1 & 1 & ms & 1 \\ 1 & 9.8 & 1 & 1 & 2 & 1 \\ 0 & 1.1 & 0 & 0 & 1 & 0 \\ 0 & 1.1 & 0 & 0 & 1 & 1 \\ 1 & 0.3 & 0 & 0 & 1 & 0 \end{bmatrix} . \qquad (3)$$

Here, it would be reasonable to replace "*ms*" by "2" since the first two instances are very similar. However, mean/mode value-based imputation methods would replace "ms" by "1" as it is the most common value for this attribute in the dataset.

One of the most powerful approaches to missing values estimation is by multiple imputation [2]. The idea is to generate multiple simulated values for each incomplete instance, and iteratively analyze datasets with each simulated value substituted in turn. The purpose is to obtain estimates that better reflect the true variability and uncertainty in the data than are done by regression. Multiple imputation methods yield multiple imputed replicate datasets each of which is analyzed in turn. The results are combined and the average is reported as the estimate. For continuous attributes, reliable estimates are obtained by combining only a few imputed datasets.

A clustering based approach for missing data imputation was considered as a local alternative to global estimation [3]. The premise was that instances could be grouped such that all the imputations in identified groups are independent from other groups. However, previous distance-based [4] clustering work was focused mainly to development of supervised clustering methods and mean/mode based imputations in these clusters. Also, prior studies were based on a strict separation for objects within clusters, such that it was assumed that there is no influence of instances in one cluster to an imputation process in other clusters.

In our DCI approach an independent cluster of similar instances with no missing values is constructed deterministically around each instance with a missing value. In contrast to a typical clustering method, we allow cluster intersections such that the same instance may be included in many clusters. DCI relies on a distance measure that considers together categorical and continuous variables and is applicable for estimation of missing values in high dimensional mixed type data.

2 Methodology

We assume that the given data consist of *M* instances with *N* attributes where *N* is a mixture of tens to hundreds of categorical and continuous attributes. The methodology section contains discussion on choice of: a measure of dissimilarity between instances in a mixed type dataset (Section 2.1); a clustering algorithm for identification of similar instances (Section 2.2); cluster-specific algorithms for imputation of missing values (Section 2.3); and imputed data evaluation along with alternative approaches (Section 2.4).

2.1 Measuring Dissimilarity between Instances in Mixed Type Data

The Minkowski distance, the Simple Matching Coefficient, the Jacquard Similarity Coefficient and other metrics could be used separately to measure the distance between instances for each type of attributes. However, such approaches are of limited applicability for mixed type data consisting of categorical and continuous attributes in the presence of many missing values [5]. Given *N* dimensional data, to measure the dissimilarity between two instances x_i and x_j of mixed type in the presence of missing values, we compute [6]:

$$\mathrm{dst}\left(x_i, x_j\right) = \left[\sum_{n=1}^{N} \frac{\left|x_{i,n} - x_{j,n}\right|}{\max\limits_{p=1\ldots P_n} x_{p,n} - \min\limits_{p=1\ldots P_n} x_{p,n}} \delta_{i,j}^{(n)}\right] \Big/ \sum_{n=1}^{N} \delta_{i,j}^{(n)}, \qquad (4)$$

$$\delta_{i,j}^{(n)} = \begin{cases} 0, & \text{if one of } x_{i,n} \text{ or } x_{j,n} \text{ is missing;} \\ 1, & \text{otherwise} \end{cases}$$

where max and min are computed over all non-missing vales of the *n*-th attribute.

2.2 Identification of Similar Instances by Clustering

To identify similar instances we employ a new clustering algorithm consists of the following steps:

1. Computing the similarity matrix (*SM*) for all instances:

$$SM = \begin{pmatrix} \infty & \mathrm{dst}(x_1, x_2) & \dots & \mathrm{dst}(x_1, x_M) \\ \mathrm{dst}(x_2, x_1) & \infty & \dots & \mathrm{dst}(x_2, x_M) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{dst}(x_M, x_1) & \mathrm{dst}(x_M, x_2) & \dots & \infty \end{pmatrix}. \quad (5)$$

2. Computing the neighborhood matrix (*NM*):

$$NM = \begin{pmatrix} nm_{1,1} & \dots & nm_{1,M} \\ \vdots & \ddots & \vdots \\ nm_{M,1} & \cdots & nm_{M,M} \end{pmatrix}, \quad (6)$$

where $nm_{i,j}$ is the number of common neighbors for instances i and j, when K nearest neighbors are counted for each of the instances.

3. For each missing value $x_{i,j}^{ms}$ compute an ordered (by ascending sort) list of all neighbor instances with no missing in j-th attribute:

$$list_{i,j} = \mathrm{sort}\left\{ \mathrm{dst}\left(x_i^{ms}, x_p\right) / nm_{i,p}, \quad p = \overline{1, P_j}; nm_{i,p} > 0 \right\}, \quad (7)$$

where x_i^{ms} denotes i-th instance with missing value in j-th attribute, and x_p denotes p-th instance with no missing in j-th attribute. Here, if two instances have the same dst/nm rate, one with less missing attributes is listed first in the list.

4. For each missing value $x_{i,j}^{ms}$ create a cluster $C_{i,j}$ by getting first R elements of $list_{i,j}$, where R is a user-specific parameter that defines a cluster size.

2.3 Cluster-Specific Methods for Imputation of Missing Values

In clusters discovered as described in Section 2.2 using similarity measure introduced in Section 2.1 we consider several possibilities for imputation of missing values based on (a) the cluster mean value of the corresponding attribute, or (b) similarity to the nearest instance with a non-missing value.

In cluster mean based imputation methods, various metrics were considered for averaging. Similarly, when relying on the nearest instance with a non-missing value for imputation several metrics were considered for identification of the nearest instances from the same cluster.

Here, we use categorical and continuos data specific imputation methods that provide a balance in terms of imputation quality and computational complexity.

Categorical data. This method estimates a missing value using the non-missing corresponding attribute in an instance from the same cluster that has the largest number of common neighbors with the imputed instance:

$$x_{i,j}^{ms} = x_{r,j}^{(C_{i,j})} : \max_{r=1\ldots R}\left\{nm_{i,r}\right\} \quad . \tag{8}$$

Continuous data. This method estimates a missing value based on all instances in the cluster that contain the corresponding attribute where each value is weighted by the *NM* between this particular instance and the imputed instance:

$$x_{i,j}^{ms} = \left[\sum_{r=1}^{R} x_{r,j}^{(C_{i,j})} nm_{i,r}\right] \Big/ \sum_{r=1}^{R} nm_{i,r} \quad . \tag{9}$$

2.4 Evaluation Measures and Alternative Imputation Methods

For evaluating imputation quality different measures were used when compare imputed categorical vs. imputed numerical data versus the corresponding true values.

The mean and absolute squared error measurements tend to be very sensitive to outliers. So, for continuous attributes and for a given tolerance τ we measured a relative prediction accuracy (RPA) defined as

$$RPA_{\tau} = \frac{n_{\tau}}{Q} \times 100\% \quad , \tag{10}$$

where n_{τ} is the number of imputed elements estimated within τ percent of accuracy from the true value of the corresponding missing value and Q is the total number of imputed values in the data. RPA is very useful in practice as an absolute precision of imputed continuous values is often not needed. A nice property of RPA measure is that it is not affected by an individual incorrect imputation (e.g. large value instead of small) that could affect considerably some statistical measures (e.g., MSE),.

For measuring imputation error in categorical attributes misclassification error (ME), or presence of incorrectly imputed values, was measured as

$$ME = \frac{s}{Q} \times 100\% \quad , \tag{11}$$

where *s* is the number of elements whose imputed values differs from originals.

As a simple imputation alternative to DCI, we used a WEKA implementation [7] of Mean and Mode Replacement (denoted here as MMR). We also compared DCI to four statistically well-founded techniques: Multiple Imputation [1], Predictive Mean Matching [8] (denoted here as PMM), Linear Regression [9], and Multilevel Regression [9] (denoted here as MLR).

The Multiple Imputation Method used for comparison and implemented in Amelia II software [10] enables to draw random simulations from the multivariate normal observed data posterior, and uses standard Expectation Maximization (EM) for finding an appropriate set of starting values for data argumentation. Multiple Imputatition begins with EM and adds an estimation of uncertainty for receiving draws from the correct posterior distribution followed by a resampling based on importance. According to King et all [10], this way is faster than traditional multiple imputation approaches, does not rely on Markov chains, produces the fully independent imputations and allows the use of about 50% more information.

Predictive Mean Matching comparison method implemented in WinMICE software [11] combines both parametric and nonparametric techniques. It imputes missing values by means of the nearest neighbors calculation. Donor where the distance is computed on the expected values of the missing variables conditional on the observed covariates, instead of directly on the values of the covariates.

Linear and Multilevel Regression models, also implemented in WinMICE, are well known statistical approaches that allow variance in imputed variables to be analyzed at multiple hierarchical levels, whereas in linear regression all effects are modeled to occur at a single level.

3 Results and Discussion

We first performed experiments on a social science dataset with mixed-type attributes to compare quality of imputation by the proposed method and alternatives in presence of various fractions of missing values. In another set of experiments mixed-type data preprocessed by various imputation methods was used for classification by several algorithms to determine practical effects of an imputation method on classification accuracy (reported in Section 3.2),.

A public domain “Adult” dataset, from the UCI Machine Learning Repository [12] was used for comparing different data imputation methods. The dataset contained a subset of records about the US population collected by the US Census Bureau. The 48,842 individuals in this database are described by 8 categorical and 6 continuous attributes (with some missing data) related to prediction of annual income. In our experiments etalon data with 46,043 instances was constructed by removing all instances from the Adult dataset with missing values. To make the dataset balanced in terms of different attribute types, two categorical attributes (“education” and “native country”) were also removed.

Eight test datasets with missing values ("holes") were constructed by randomly hiding 0.2%, 0.5%, 1.1%, 1.8%, 5.4%, 10.9%, 16.3% and 32.6% of data elements in both categorical and continuous attributes of the etalon data. Each test database was fully independent from others, which means that places of "holes" were independent.

3.1 Evaluation of Imputation Quality on Mixed Type Data

The DCI and other imputation algorithms described in section 2.4 were compared using eight datasets with different fraction of introduced missing values. Imputed values were compared to the true values in Adult dataset. The misclassification rate of imputation is summarized in Table 1..

Table 1. Misclassification error (ME) for categorical attributes.

Imputation Methods	ME for different fractions of missing values							
	0.2%	0.5%	1.1%	1.8%	5.4%	10.9%	16.3%	32.6%
DCI	33.98	30.81	32.46	29.97	28.96	28.53	29.36	34.36
MMR	45.51	62.01	46.32	43.85	45.56	45.25	45.25	45.28
PMM	65.82	62.21	64.09	63.91	63.18	64.11	64.38	64.80
Linear Regression	71.88	69.66	71.22	71.04	71.04	71.55	71.65	71.92
Multiple Imputation	53.13	50.90	50.96	50.17	51.93	52.21	52.66	54.48
MLR	70.70	70.33	72.08	70.23	71.18	71.54	71.38	71.82

Misclassification error analysis for imputation of categorical attributes (Table 1) revealed that for all fractions of missing values DCI was a lot more accurate from alternative five imputation methods (1.3-1.7 times more accurate than the best of the remaining methods). Mean Mode Replacement approach was the second most accurate imputation method. The results of the remaining imputation methods had more than 50% imputation error.

Table 2. Relative prediction accuracy (RPA) with 5% tolerance for continuous attributes.

Imputation Methods	RPA (5%) for different fractions of missing values							
	0.2%	0.5%	1.1%	1.8%	5.4%	10.9%	16.3%	32.6%
DCI	33.81	28.32	28.08	31.19	29.51	30.29	30.24	28.32
MMR	3.69	4.94	1.37	5.45	1.21	1.49	1.41	5.47
PMM	18.65	20.91	20.04	20.20	18.72	19.62	19.43	19.37
Linear Regression	3.69	4.48	4.42	4.50	4.21	4.35	4.39	4.22
Multiple Imputation	5.53	11.76	3.89	4.74	4.59	4.70	4.58	4.45
MLR	3.69	4.28	4.62	3.95	4.18	4.41	4.38	4.27

The Relative Prediction Accuracy of DCI for imputation of continuous attributes (Tables 2-4) was also much better from alternative imputation methods. Here, Predictive Mean Matching was the second most accurate method. For 5% tolerance DCI provided 1.35-1.81 times better accuracy from PMM and was 6-9 times other alternatives (Table 2). Comparing with other methods DCI presented overwhelming 6-9 times better accuracy except for the dataset with 0.5% fraction of missing values.

Table 3. Relative prediction accuracy (RPA) with 10% tolerance for continuous attributes.

Imputation Methods	RPA (10%) for different fractions of missing values							
	0.2%	0.5%	1.1%	1.8%	5.4%	10.9%	16.3%	32.6%
DCI	38.73	35.40	35.59	37.39	36.68	37.22	37.22	35.61
MMR	10.25	11.76	12.03	11.91	11.48	11.62	11.56	11.99
PMM	25.6	29.9	29.1	28.7	27.5	28.5	28.2	28.2

	1	9	9	5	9	4	0	5
Linear Regression	10.66	13.56	13.61	13.01	13.09	13.30	13.24	13.14
Multiple Imputation	13.32	20.44	13.40	13.48	13.30	13.34	13.35	13.27
MLR	10.66	12.89	13.94	13.01	13.13	13.18	13.26	13.17

Table 4. Relative prediction accuracy (RPA) within 15% tolerance for continuous attributes.

Imputation Methods	RPA (15%) for different fractions of missing values							
	0.2%	0.5%	1.1%	1.8%	5.4%	10.9%	16.3%	32.6%
DCI	42.01	40.15	40.01	41.84	41.53	41.96	42.02	40.51
MMR	15.57	17.57	17.49	17.37	17.39	17.42	17.39	17.83
PMM	30.33	34.74	33.91	33.45	32.62	33.51	33.13	33.25
Linear Regression	15.37	18.30	18.36	17.73	18.12	18.26	18.18	18.13
Multiple Imputation	17.01	25.52	18.06	17.83	18.21	18.10	18.25	18.12
MLR	14.96	18.10	18.53	17.51	18.19	18.16	18.24	18.17

Experiments with double and triple tolerance for estimation error of 10% and 15% (Tables 3 and 4) resulted in reduced difference in accuracy between imputation methods. However, even for larger tolerance DCI was still 20-50% more accurate (in relative difference) than the second best PMM method.

3.2 Effect of an Imputation Method on Classification Accuracy for Mixed Type Data

The next stage of our experiments was devoted to practical comparison of how well different imputation techniques would suit for real life classification tasks. The idea was to explore a scenario where clean mixed type data was used for training a classification model while it was applied to real data with various fractions of missing values. For this purpose we built several kind of classifiers by training them on the first 16,043 subjects from etalon Adult database where for each instance all 12 attributes were available. For a test subject drawn from the remaining 30,000 instances the task was to predict if he/she makes over 50,000 U.S. dollars a year where a fraction of variables was missing at random. Different fractions of missing values were considered and preprocessing was achieved by 6 imputation methods described in Section 2. As a measure of accuracy, the percent of Incorrectly Classified Instances (ICI) was calculated.

As a classification method we applied three models implemented in WEKA: NB-Tree [13], Random Subset Selection [14] and Multilayer Perceptron [15]. NB-Tree was used as one of the best classification methods for "Adult" database according to [12]. Random Subset Selection and Multilayer Perceptron were used as alternative solutions that showed good speed and classification accuracy in other domains, respectively. The experimental results for all classifiers are reported in Tables 5-7.

Table 5. Incorrectly Classified Instances (ICI) of NB-Tree classification model applied to datasets with various fraction of imputed values.

Imputation Methods	ICI of NB-Tree for different fractions of missing values							
	0.2%	0.5%	1.1%	1.8%	5.4%	10.9%	16.3%	32.6%
DCI	13.88	13.90	13.99	14.00	14.25	14.64	15.16	16.43
MMR	13.90	14.01	14.06	14.11	14.87	15.84	17.00	20.43
PMM	13.88	13.92	14.13	14.14	14.96	15.53	16.37	19.01
Linear Regression	13.90	14.00	14.31	14.39	15.73	17.11	18.63	23.21
Multiple Imputation	13.88	13.94	14.26	14.35	15.45	16.92	18.57	22.87

MLR	13.90	13.96	14.24	14.36	15.72	17.12	18.65	23.20
Etalon dataset	13.86	13.86	13.86	13.86	13.86	13.86	13.86	13.86

All imputation methods had very similar accuracy for 0.2-1.8% fractions of missing values (Table 5). The difference was more substantial when more than 5% of missing values were imputed. Though DCI provided NB-Tree classifier with the best imputed datasets within all fractions of missing values, its advantage was the most evident for the largest fraction of missing values (32.6%) where it had 14-22% lesser error than alternative methods.

Table 6. Incorrectly Classified Instances (ICI) of Random Subspace Selection classification model applied to datasets with various fraction of imputed values.

Imputation Methods	ICI of Random Subset for different fractions of missing values							
	0.2%	0.5%	1.1%	1.8%	5.4%	10.9%	16.3%	32.6%
DCI	15.12	15.11	15.14	15.17	15.20	15.03	15.35	15.24
MMR	15.12	15.13	15.12	15.24	15.47	15.66	16.23	18.40
PMM	15.15	15.10	15.21	15.28	15.60	15.80	15.94	16.95
Linear Regression	15.16	15.12	15.24	15.34	15.85	16.34	16.61	17.93
Multiple Imputation	15.11	15.13	15.19	15.23	15.52	15.80	16.33	17.41
MLR	15.12	15.12	15.22	15.28	15.70	16.16	16.68	18.20
Etalon dataset	15.11	15.11	15.11	15.11	15.11	15.11	15.11	15.11

Random Subset classifier exposes (Table 6) same qualities of imputation methods as NB-Tree classifier. However, it showed better overall tolerance for increasing fraction of missing values . Once again, DCI outperformed alternative approaches on the largest fractions of missing values for 11-20%.

Table 7. Incorrectly Classified Instances (ICI) of Multilayer Perceptron classification model applied to datasets with various fraction of imputed values.

Imputation Methods	ICI of Multilayer Perceptron for different fractions of missing values							
	0.2%	0.5%	1.1%	1.8%	5.4%	10.9 %	16.3 %	32.6 %
DCI	15.46	15.44	15.48	15.44	15.36	15.30	15.48	15.35
MMR	15.51	15.50	15.56	15.61	15.89	16.42	16.98	19.22
PMM	15.49	15.46	15.68	15.68	16.23	16.81	17.31	19.18
Linear Regression	15.51	15.52	15.80	15.87	16.80	18.02	19.02	22.64
Multiple Imputation	15.48	15.53	15.73	15.83	16.64	17.70	18.82	22.53
MLR	15.51	15.55	15.73	15.79	16.78	17.80	19.15	22.54
Etalon dataset	15.47	15.47	15.47	15.47	15.47	15.47	15.47	15.47

Neural Network based classifier, represented by a 3-layer perceptron showed (Table 7) similar characteristics to NB-Tree and Random Subspace Selection. DCI imputation resulted in more accurate classification in all datasets with a large fraction of missing values. For 0.2%, 1.1%, 5.4% 10.9%, and 32.6% preprocessed datasets it showed somewhat better accuracy than etalon data, which may be accounted for Multilayer Perceptron nature.

To address a considerable misbalance for target variable in the “Adult” dataset (34,621 subjects in one class vs. 11,422 in another) we also measured Kappa coefficient and F-score for three classification models performed on 32.6% of missing values imputed by six methods (Table 8).

Table 8. Kappa coefficient and F-score of NB-Tree, Random Subspace Selection, and Multilayer Perceptron classification models applied to datasets with 32.6% of missing values imputed by 6 methods and to complete data without missing values.

Imputation Methods	NB-Tree		Random Subset		Multilayer Perceptron	
	κ	F	κ	F	κ	F
DCI	0.518	0.826	0.543	0.837	0.535	0.834
MMR	0.416	0.788	0.493	0.813	0.447	0.800
PMM	0.470	0.806	0.472	0.813	0.440	0.798
Linear Regression	0.372	0.767	0.442	0.802	0.374	0.770
Multiple Imputation	0.398	0.773	0.477	0.813	0.384	0.772
MLR	0.372	0.766	0.432	0.799	0.376	0.770
Etalon dataset	0.611	0.858	0.548	0.838	0.532	0.833

The obtained resutls clearly suggest that DCI based pre-processing resulting in the nearest accuracy to etalon dataset in the sense of both Kappa coefficient [16] and F-score statistics [17]. We also observe that our resutls on imputed data confirms previous findings obtained on full data that NB-Tree based classifier is the best choice for classification of Adult data. However, we also observe that the most stable results in terms of accuracy were obtained by Random Subset classifier.

4 Conclusion

Data imputation to replace missing values is often an important preprocessing step in data analysis. This study identified some limitation of commonly used and of known statistical methods when applied to mixed type data with a large fraction of missing values. In our approach, the main idea was to make all replacements independently for data within clusters created around each missing value. This is theoretically reasonable and is useful for a practical implementation. Our experiments on a social science mixed type data provide evidence that the proposed data imputation methodology is more accurate than the most commonly used alternatives and is effective when a large fraction of data is missing.

While the computational complexity of the proposed imputation method of $O(M^3 \log M)$ is a limiting factor in large scale applications, many possibilities for improvements remain. For example, cluster-specific multiple imputation techniques based on DCI idea could be developed. Also, specialized algorithms for defining optimal size of specific clusters may be created. Finally, organizing data to KD-trees may improve the overall matrix processing speed.

Acknowledgments: This study was supported in part by Award No. 2006-IJ-CX-0022 awarded by the National Institute of Justice, Office of Justice Programs, US Department of Justice. The opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Department of Justice. Funding support from Russian Ministry of Science and Education is also acknowledged.

Authors thank the Department of Human Services (DHS), Philadelphia, and the Crime and Justice Research Center at Temple University for creating ProDES [18] and granting us access to the dataset. During this study we found a real-life data imputation problem that initiated DCI algorithm development.

We also thank Dr. Alan Izenman and Dr. Jeremy Mennis for extremely valuable comments on an early version of the methods and the results reported in this manuscript. Finally, we thank Pavel Karpukhin who rewrote all Matlab code Matlab and made it possible to do most of our experiments.

References

Little, R.J.A., Rubin, D.B.: Statistical analysis with missing data. John Wiley & Sons, New York (1987)
Schafer, J.L.: Multiple imputation: a primer. Statistical Methods in Medical Research. Vol. 8, 1, 3--15 (1999)
Fujikawa, Y., Ho, T.B.: Cluster-based algorithms for dealing with missing values. In: Knowledge Discovery and Data Mining Conference, pp. 549--554. Springer, Berlin (2002)
Mantaras, R.L.: A distance-based attribute selection measure for decision tree induction. Machine Learning. Vol. 6, 81--92 (1991)
Gan, G., Ma, C., Wu, J.: Data Clustering: Theory, Algorithms, and Applications. SIAM Press, Philadelphia (2007)
Wishart, D.: K-means clustering with outlier detection, mixed variables and missing values. In: Schwaiger M., Opitz O. (eds.) Exploratory Data Analysis in Empirical Research, pp. 216--226. Springer, New York (2003)

Witten, I.H., Frank, E.: Data Mining: Practical machine learning tools and techniques, 2nd Edition. Morgan Kaufmann, San Francisco (2005)

Landerman, L.R., Land, K.C., Pieper, C.F.: An Empirical Evaluation of the Predictive Mean Matching Method for Imputing Missing Values. Sociological Methods & Research, Vol. 26, 1, 3--33 (1997)

Gelman, A., Hill, J.: Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, Cambridge (2006)

King, G., Honaker, J., Joseph, A., Scheve, K.: Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. American Political Science Review. Vol. 95, 1, 49--69 (2001)

Oudshoorn, C.G.M., Buuren, V.S., Rijckevorsel V. Flexible Multiple Imputation by Chained Equations of the AVO-95 Survey. In: TNO Prevention and Health. Report PG/VGZ/99.045 (1999)

Asuncion, A., Newman, D.J.: UCI Machine Learning Repository. University of California, Irvine, http://www.ics.uci.edu/~mlearn/MLRepository.html

Kohavi, R.: Scaling Up the Accuracy of Naive-Bayes Classifiers: A Decision-Tree Hybrid. Proc. In: 2-nd Int. Conf. on Knowledge Discovery and Data Mining, pp. 202--207. AAAI Press, Portland (1996)

Ho, T.K.: The Random Subspace Method for Constructing Decision Forests. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 20, 8, 832--844 (1998)

Haykin, S.: Neural Networks: A Comprehensive Foundation, 2-nd ed. Prentice Hall (1998)

Gwet, K.: Statistical Tables for Inter-Rater Agreement. StatAxis Publishing, Gaithersburg (2001)

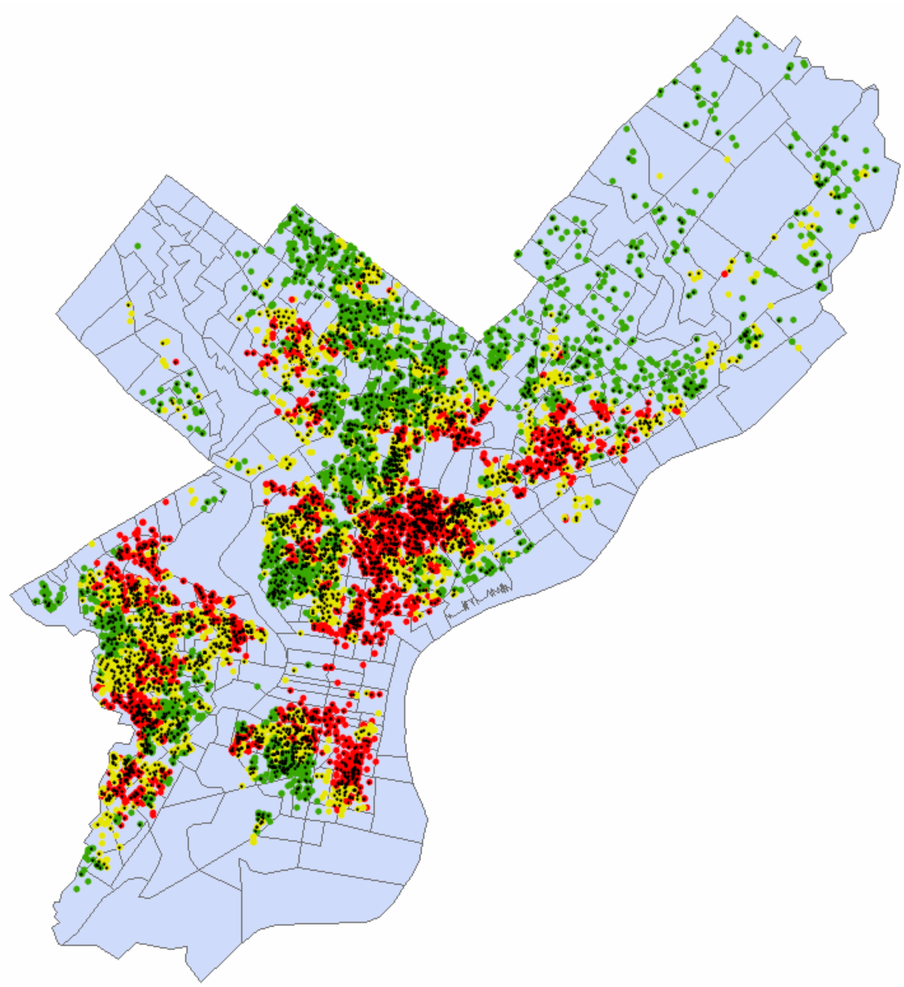
Van Rijsbergen, C.J.: Information Retrieval. 2-nd edition. Butterworth, London (1979)

Crime and Justice Research Center, Temple University, http://www.temple.edu/prodes/

APPENDIX C: ADDITIONAL MAPS

Hot Spot Analysis: Classification of Recidivism Ratios

Map C.1



Red=High

Yellow=Medium

Green= Low

Black=Instances with Recidivism

Map C.2

Recidivism Ratios, Males only

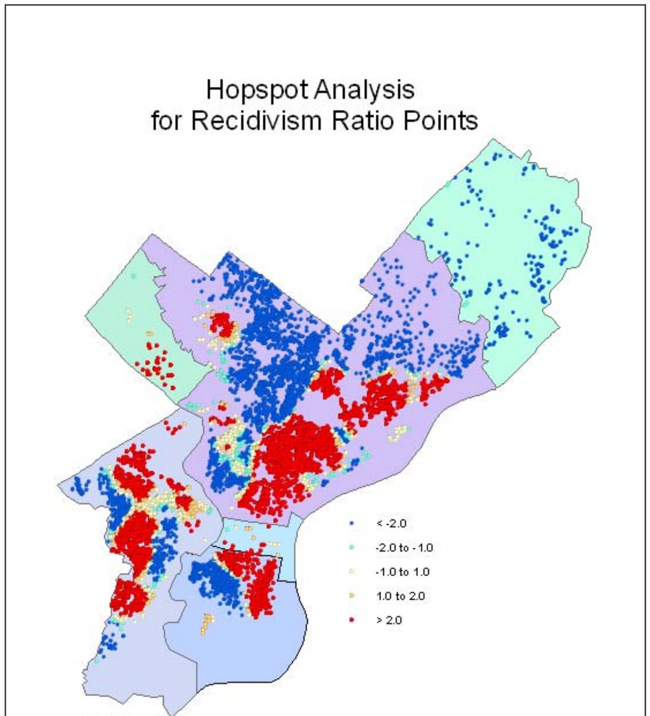


Figure Ganypet:
Hotspot analysis was performed for the recidivism ratio points for males only in the a region merged from six small regions

Map C.3

New Offenses Against Persons, Males only

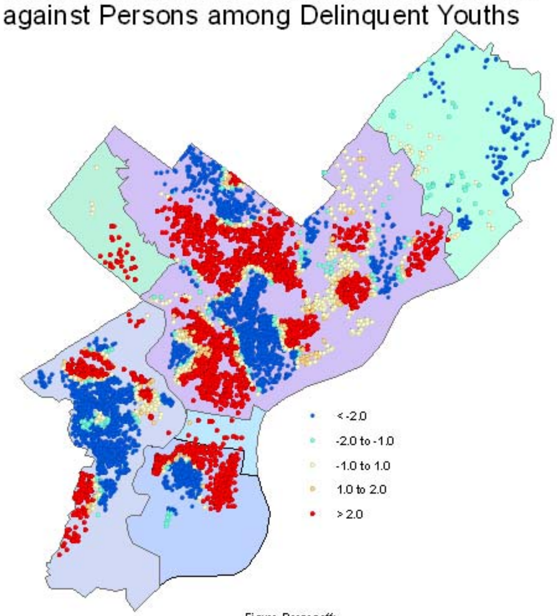


Figure Personoff:
Hotspot analysis is performed on the ratio for males only in a region merged from six small regions