ECONOMIC FACTORS IN THE PREDICTION OF RECIDIVISM

by

John Kenneth Krantz

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Economics

The University of Utah

May 2010

ABSTRACT

This dissertation examines the impact of economic variables on recidivism and discusses the implications for criminal justice policy. The definition, measurement, and policy uses of recidivism rates are discussed and the importance of reducing recidivism is emphasized through a discussion of the costs of crime and, in particular, criminal justice spending in Utah. International comparisons of recidivism, incarceration, and crime rates are made between the United States and 16 countries in order to evaluate where the United States stands relative to other countries.

Past recidivism studies have tended to focus on sociological, or noneconomic variables, instead of economic variables. While some noneconomic variables may have predictive value, they are ultimately unimportant for policy because they are unchangeable. In contrast, economic variables reflect behavioral responses to incentives and disincentives. Policies designed to change the structure of incentives that individuals face can influence their behavior. Thus, if economic variables are important determinants of recidivism, these variables should be of central concern to the formation of criminal justice policy.

An analysis of parolees in Utah confirms that economic variables are strong predictors of recidivism. Recidivism is modeled from the Utah parolee data using the techniques of classical linear regression, Bayesian Model Averaging, and the Classification and Regression Tree methodology. In all of the models, the economic variables of employment status, the payment of restitution, and the payment of child support are among the strongest predictors of recidivism.

The results of the Utah parolee data analysis have important implications for criminal justice policy. The costs and benefits of policies regarding employment, restitution payments, and child support payments are examined using the Utah parolee data. The results suggest that by improving the ability of released prisoners

to find employment, eliminating restitution payments, and alleviating the burden of child support payments corrections costs can be reduced significantly. The importance of these results for criminal justice policy, both in terms of lower corrections costs and the successful reintegration of offenders into society, underscores the need for further research. This study concludes with a consideration of several topics for future research suggested by the Utah parolee data.

For Tisha and Mimi

CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Corrections spending is placing an ever-growing burden on state and local budgets across the United States. As an historical average, approximately 90% of corrections costs are covered by state and local expenditures with federal expenditures covering the remaining 10%. Table 1.1 lists the percentage change in expenditures for various budgetary categories for all state and local expenditures in the United States from 1977 to 2007 (U.S. Census Bureau, 1980, 2009). Total direct general expenditure represents the growth of total state and local spending from the general funds and is used as a basis of comparison for examining the relative growth of the other expenditure categories. Not all of the categories of state and local expenditure are included in Table 1.1 because the limited decomposition of the available 1977 data created problems of comparability. Nevertheless, the categories listed represent more than 80% of all state and local direct general expenditures for 2007.

While several categories of expenditure have increased faster than the rate of total expenditure increase, none has grown faster than corrections spending. Corrections spending has increased at nearly twice the rate of the increase in state and local total direct general expenditure. Most of the increase in corrections spending can be explained by one fact alone: The total incarcerations per 100,000 citizens in the United States increased from 132 in 1977 to 762 in 2007, an increase of 477% (Sabol & Couture, 2008; U.S. Census Bureau, 1980).

With corrections spending increasing at such a rapid pace, its burden is felt primarily through the squeeze it places on other important budgetary categories. State governments cover roughly 60% of all corrections costs, implying that the

Table 1.1. U.S. State and Local Expenditures 1977-2007

Expenditure Category	Percent Change
Total Direct General Expenditure	722
Corrections	1369
Housing and Community Development	1256
Health	1180
Public Welfare	993
Fire Protection	735
Police Protection	733
Interest on General Debt	728
Education	656
Higher Education	687
Elementary and Secondary Education	648
Hospitals	576
Highways	528

burden rests mostly upon state budgets. State governments are also responsible for providing the majority of higher education funding. The impact of corrections spending on state budgets can be appreciated by considering the ratio of corrections spending to higher education spending over time. From 1987 to 2007, the ratio of corrections spending to higher education spending increased from .32 to .6 for all state government general expenditure (Pew Center on the States, 2008). If the percentage shares of some budgetary expenditure categories increase, other expenditure categories must necessarily decrease. Referring to Table 1.1, expenditures directed toward education, hospitals, and highways can be viewed as having been crowded out by all of those expenditure categories that increased at a rate higher than 722%. Expanding by 1,369% over the 30 years from 1977 to 2007, corrections spending has played a significant role in crowding out educational spending as well as other spending categories. Corrections spending has historically accounted for only a relatively small share of state budgetary spending, amounting to only 1.7% of all state direct general expenditures in 1977. However, given the rate at which corrections spending has increased, it is becoming an increasingly conspicuous component of state budgets. According to the most current figures available, state governments spent on average 6.8% of their general funds on corrections in 2007 and 11% of all state employees worked for corrections in 2006 (Pew Center on the States, 2008).

Recent events in California illustrate the severity of the problems associated with an ever-increasing incarceration rate and its concomitant increase in corrections spending. In 2007, California spent $8.8 billion on corrections, by far the largest amount spent by any state in the country (Pew Center on the States, 2008). With the economic recession worsening California's budget deficit, spending cuts for most programs appear inevitable. Corrections spending, however, will most likely not experience any spending cuts because Republican lawmakers along with some Democrats "have their eyes on higher office and don't want to appear soft on crime" (California prison spending, 2009, ¶3). Higher education is, in essence, being crowded out by corrections spending as "California now spends more incarcerating 167,000 adults than it does to educate 226,000 students in its 10-campus University of California system" (¶7). The incarceration rate in California is actually increasing faster than the expansion of corrections spending, which has led to overcrowding in prisons. On August 8, 2009, 175 people were injured in a 4-hour riot at a prison in Chino, California that held 5,900 men but was designed for 3,000, and the cause was attributed to overcrowding (California prison riot, 2009). The riot forced lawmakers to pass measures designed to release up to 27,000 inmates throughout the state, including elderly, medically-disabled, nonviolent, and other low-risk offenders (California Senate, 2009).

While the problems associated with an ever-growing prison population and the accompanying corrections costs have not been as extreme in Utah as in California, they are, nevertheless, growing concerns. Utah's total incarceration rate for 2005 was relatively low at 466 per 100,000 Utah citizens, but Utah's prison population continued to grow by 1.6% during 2007 (Pew Center on the States, 2008). As a consequence, the increasing share of the state budget devoted to corrections spending has necessarily crowded out other expenditure categories. Corrections spending accounted for only 3.8% of the state general funds for the fiscal year

1984-85 (State of Utah Governor's Office of Budget and Planning, 1988). However, 7.1% of the general funds for the fiscal year 2010 have been appropriated for corrections (State of Utah Governor's Office of Budget and Planning, 2009). The crowding out effect of the increase in corrections spending can be illustrated by considering the ratio of corrections expenditure to higher education expenditure over time. From the general funds for the state government of Utah, the ratio of corrections to higher education spending increased from .23 in 1987 to .41 in 2007, and is projected to increase to .46 in the 2010 budget (Pew Center on the States, 2009; State of Utah Governor's Office of Budget and Planning, 2009).

The preceding discussion was intended to give substance to the problems of the rising incarceration rate and increasing corrections costs. Many potential solutions have been offered to deal with the high incarceration rate. Among the most frequently discussed solutions are eliminating mandatory sentencing laws, seeking alternative punishments for low-risk offenders, and giving parole boards more flexibility with respect to early releases and the decision to reincarcerate individuals for technical parole violations. This dissertation, however, focuses solely upon one solution to address the problem of the high level of incarcerations: reducing the recidivism rate.

Specifically, this study centers upon an examination of economic factors that influence the recidivism rate. From a behavioral perspective, certain policies designed to reduce recidivism are better solutions than those of modifying sentencing laws, releasing low-risk offenders, and giving parole boards more flexibility. The latter policies may influence offender behavior through the deterrence effect, but this influence is indirect and probably insignificant. Moreover, if the effect of these latter policies is not altogether negligible, it likely reduces the deterrence effect. On the other hand, policies that effectively reduce recidivism by modifying incentives and disincentives must be deemed superior because they directly eliminate the criminal behavior rather than merely changing the standard by which a behavior is deemed punishable by incarceration. In the development of criminal justice policies to reduce recidivism, economic factors would appear to be the best candidates for modification because their influence on behavior is direct and predictable. Moreover, economic factors are easily modified.

The potential usefulness of results derived from investigations into the impact of economic factors on recidivism would seem to imply that research on the subject must be relatively abundant. On the contrary, such studies are rather scarce. Nationwide statistics describing the economic conditions of those released from prison are extremely limited or, in the case of most variables of interest, nonexistent. For example, a recent study examining the repayment of debt by offenders that was funded by the U.S. Department of Justice could only cite restitution and child support statistics from small regions based on personal communications rather than from published research or carefully collected official data (McLean & Thompson, 2007). The Pew Center on the States (2008) also affirms that economic statistics related to parolees and probationers are scarce. It is likely that the explanation for the lack of research into economic factors and their influence on offender behavior rests upon a combination of economic, political, and legal reasons. In any event, regardless of the end results of the analysis, this study will be of some importance simply as a contribution to the currently small quantity of research on the subject.

The purpose and structure of this dissertation are described in the following overview. At the core of this study is a statistical analysis of factors influencing recidivism among parolees in Utah. Chapters 2 through 4 cover issues that are preliminary to the analysis of recidivism. Chapter 2 contains a discussion of the various ways in which recidivism is defined and measured. Recidivism rates may be appropriate for measuring the effectiveness of some programs and policies, but not others. A few issues concerning the proper use and misuse of recidivism rates are also addressed. In order to develop a perspective on the magnitudes of the recidivism rates for the United States and Utah, Chapter 3 is devoted to the international comparisons of recidivism rates between the U.S. and 16 other relatively similar countries. In addition, international statistics are used to examine the relationship between the crime and incarceration rates and the recidivism and crime rates. Chapter 4 reviews the literature on predictors of recidivism. The content of Chapter 4 is organized around the distinction between economic and noneconomic predictors of recidivism in order to determine the relative importance of economic predictors of recidivism in past studies.

The analysis of recidivism among Utah parolees is undertaken in Chapter 5. After discussing the survey and the descriptive statistics of the Utah parolees, attention focuses on three types of models used to analyze the data. Emphasis is placed on the results of two Bayesian modeling techniques: Bayesian Model Averaging and Classification and Regression Trees. In addition, a classical linear probability model is estimated for the purpose of comparisons. The three models are compared with respect to variable selection within each model, the prediction of the recidivism rate, and the overall model fit of the data. The results demonstrate that economic variables are important in the prediction of recidivism.

Chapter 6 concludes with discussions of the policy implications of the analysis and proposals for future recidivism research with respect to economic factors. Using the Utah parolee data along with cost information, criminal justice policies that affect the economic incentives and disincentives of parolees are critically analyzed. Even though certain policy recommendations will appear to be beneficial to society as a whole, the nature of the political system makes such policy changes difficult to implement. The chapter closes with proposals for future research into economic variables that may potentially influence recidivism. With the importance that economic factors play in recidivism and the potential for sizable reductions in corrections spending implied by reducing recidivism, there is an urgent need for future research into other economic factors that affect parolees and probationers.

CHAPTER 2

RECIDIVISM RATES: DEFINITIONS AND POLICY

While the meaning of the term *recidivism* is intuitive, practical difficulties in the measurement of recidivism have led to many different definitions of the concept. Each definition has its advantages and disadvantages. The most commonly used measures and their associated problems are discussed in the first section.

In addition to there being many definitions of recidivism, there are also many ways in which a recidivism rate can be used. Changes in the recidivism rate are most frequently used to measure the effectiveness of corrections programs, but they are often used to measure the effectiveness of various forms of criminal legislation as well. Criticism, however, has been raised concerning what recidivism rates can, in fact, effectively measure and a few comments are directed toward this issue. A contentious topic associated with the policy uses of recidivism rates is the so-called *what works* debate, an event that produced policy conclusions that were extreme and likely unjustified. The second section closes with a brief overview of the debate and its significance for this current study.

2.1 Definitions and measurements of recidivism

Recidivism derives from the Latin term *recidere*, which means *to fall back*. As it is used within a criminological sense, recidivism refers to the relapse into criminal behavior after receiving punishment for previous criminal activities. Recidivism is most frequently measured as a rate defined as the number of recidivists divided by the total number of released offenders in an observed cohort. The measurement of a recidivism rate only requires the specification of a time period over which the offenders are observed, referred to as the follow-up period, and a set of criteria for classifying offenders as recidivists at the end of the time period. The difficulty arises when deciding upon the criteria used to determine whether an offender has committed a new crime. The different measures of recidivism lead to differences in the estimation of the recidivism rate and differences in the likelihood of classification errors. The subjective evaluation of the magnitude of recidivism can depend upon the particular definition of recidivism used. For example, law enforcement officials might believe that the recidivism rate is high when based upon the rearrest rate, while at the same time corrections officials might believe that the recidivism rate is low when based upon the reincarceration rate (Blumstein & Larson, 1971). The logical starting point for a discussion of the various definitions and measures of recidivism is a characterization of the true recidivism rate.

The true or actual recidivism rate represents the ideal measurement of recidivism, an ideal that most likely can never be realized. In practice, the measurement of recidivism is based only on contact with the criminal justice system and there will always be some amount of criminal activity such that the perpetrator goes unidentified, the crime goes unreported, or the crime goes undetected. The measurement of the true recidivism rate would seem to require offenders to self-report their crimes, a task that offenders will be reluctant to perform. As a result, Blumstein and Larson (1971) conclude that all practical measures of recidivism necessarily underestimate the true level of recidivism.

Turning to the practical measures of recidivism, the rearrest rate tends to produce the highest rate among all of the practical measures and is most likely the closest approximation to the true recidivism rate. The rearrest rate will be below the true recidivism rate because not all perpetrators are arrested and not all crimes are reported. Even though it is below the true recidivism rate, it may still incorrectly classify recidivists. The errors associated with incorrectly classifying recidivists can be expressed in terms of type I and type II errors. The occurrence of a type I error is when an individual is classified as a recidivist, but in fact is not one, while the occurrence of a type II error is when an individual is classified as a nonrecidivist, but actually is one. Maltz (2001) notes that the rearrest rate may produce type I errors for several reasons. An individual can be arrested only on the basis of probable cause and as a less-than-perfect indicator of criminality it may lead to the arrest of the innocent. Those with prior arrest records are likely to be rearrested when a major crime has been committed and later, when found innocent, released. Moreover, if police departments are evaluated on the basis of crimes cleared through arrests, there may be an incentive to produce a high number of arrests. The police may also arrest individuals for the purpose of harassment. All of these reasons lead to the conclusion that the rearrest rate likely overestimates the recidivism rate for those having direct contact with the criminal justice system.

The reconviction rate produces a recidivism rate that is necessarily lower than the rearrest rate because arrest is always prior to conviction. The reconviction of an offender requires satisfying the standard of proof beyond a reasonable doubt, which is a higher standard than probable cause. While this reduces the probability of type I errors, it increases the probability of type II errors. In addition to the errors resulting from the higher standard of proof, type II error may result from charges being dropped in exchange for testimony, cases being dismissed due to heavy workloads in the court system, and plea bargaining leading to different charges (Maltz, 2001). For these reasons, it appears likely that the reconviction rate underestimates the recidivism rate for those having direct contact with the criminal justice system. Maltz believes that the type II errors associated with the reconviction rate are larger than the type I errors associated with the rearrest rate and, therefore, finds the rearrest rate more accurate.

The remaining measures of recidivism mentioned here are all based on returning to prison. There are several variations in how returns to prison are measured as instances of recidivism. The term *reincarceration rate* is generally used to denote the rate of returns to prison that occur specifically as the result of a conviction for a new crime (Beck & Shipley, 1989; Langan & Levin, 2002). The term *revocation rate* is used to denote the rate at which parolees and probationers return to prison for technical violations. A point of confusion in the literature is the reference to a reincarceration rate as the sum of returns for new convictions and technical violations. There is no widely accepted term that specifically denotes returns for either new convictions or technical violations. Within this dissertation, the convention is adopted that the term *returns to prison* refers to the sum of returns for new convictions and technical violations. The reincarceration rate, used in its technical sense, produces the lowest recidivism rate and, therefore, represents the largest underestimation of the true level of recidivism. The reincarceration rate likely produces the smallest number of type I errors, but type II errors will probably be even greater than for the reconviction rate. This can result from convictions that lead to punishments not involving imprisonment. Returns to prison may include even more type II errors than the reincarceration rate due to the inclusion of technical violations. Even though technical violations do represent some form of disobedience, it is not clear that all types of technical violations actually represent relapses into criminal behavior. For example, the consumption of alcohol within prescribed circumstances is legal for all adults meeting a minimum age requirement, but it may constitute a technical violation for a parolee in all circumstances.

There are other other issues that can lead to errors in the measurement of the recidivism rate, only two of which are mentioned here. An offender may be rearrested in a state other than the one in which the offender was previously incarcerated. This out-of-state rearrest information is not always available to those conducting a recidivism study. Another problem is absconsion. In some cases, absconsion is treated as an instance of recidivism, particularly if the individual is a parolee who stops reporting and an arrest warrant is issued. However, it is not clear whether a relapse into crime has necessarily occurred. Even though the numbers of absconsions and out-of-state rearrests are small, they do lead to additional errors in the measurement of recidivism.

The measure of recidivism used in the analysis of the Utah parolee data is returns to prison. The only information that was available for measuring recidivism in this study was corrections data, which precluded the use of either rearrest or reconviction data. The decision to include both technical violations and new commitments in the measurement of recidivism was based on the connection between the economic variables and technical violations. The conditions of parole usually stipulate that the offender must find employment, make restitution payments, pay various fees, and so forth. Failure to meet these conditions can result in the revocation of parole. Thus, including new commitments and technical violations in the measure of recidivism allows for a wider, more complete measure of the effects of economic factors upon returning to prison.

2.2 The uses of recidivism rates

To understand the role of recidivism rates in the formation of criminal justice policy, a brief discussion of the reasons for the imprisonment of criminal offenders is in order. The effect of a stay in prison upon a criminal offender is of particular interest to criminal justice policy and the effect is often measured in terms of a change in the recidivism rate. In a study of the rehabilitation of criminal offenders, Sechrest, White, and Brown (1979) list the following seven justifications for the imprisonment of offenders: to deter the offender from crime in the future, to deter others from crime in the future, to incapacitate the offender from crime for a period of time, to forestall personal vengeance, to exact retribution, to educate society, and to rehabilitate the offender. In some of the cases above, the relationship to recidivism is clear (e.g., deterrence); in other cases, the relationship is not at all clear (e.g., retribution).

According to Maltz (2001), only the offender-related goals of special deterrence, incapacitation, and rehabilitation have a valid relationship to the recidivism rate. Special deterrence (i.e., the deterrence of a particular offender from future criminal activity) can be achieved through a prison sentence if it either scares the offender from committing future crimes or at least convinces the offender that the cost of crime outweighs the benefits. In either case, the effect can be measured by the recidivism rate. The measurement of the incapacitation effect of incarcerating an offender requires the use of a recidivism rate to form an estimation of how much crime would have been committed if the offender were not imprisoned. By far the most prevalent use of the recidivism rate is in evaluating programs designed to rehabilitate offenders. The recidivism rate needs to be compared only before and after the implementation of a particular rehabilitation program in order to measure its effectiveness.

Maltz (2001) considers all three of these uses of recidivism rates valid, but he is particularly critical of using a recidivism rate to measure rehabilitation. The use of programs designed to rehabilitate prisoners depends on the assumptions that the offender is in need of correction rather than society, the offender can be corrected by way of some specific program, and the correction of the problem will lead to lower criminality. Recidivism is an inherently negative criterion for the purpose of evaluation because it only identifies failures. Consequently, no attention is paid to the successes resulting from rehabilitation programs. It is certainly possible that a rehabilitation program can be highly successful in correcting particular problems among prisoners, but that no decrease in criminality results. Strictly speaking, a recidivism rate cannot be used to evaluate whether a rehabilitation program has corrected a particular problem, but it can indicate only whether criminality has decreased. Regarding the other four reasons for imprisonment mentioned above, Maltz states that recidivism rates are not appropriate for measuring changes in these society-related goals.

Recidivism rates have historically played an important role in the formation of criminal justice policy. However, poor design and control of statistical analyses have produced what might be viewed as equally poor policy decisions. The application of criminal justice policy based on recidivism rates is exemplified through two famous examples, both of which led to rather extreme policy decisions. The first example concerns a meta-analysis that sought to determine what types of rehabilitation programs led to reductions in the recidivism rate. In their study *The Effectiveness of Correctional Treatments*, Lipton, Martinson, and Wilks (as cited in Maltz, 2001) reexamined 231 studies of prison rehabilitation programs that were conducted from 1945 to 1967. The results of their analysis were widely disseminated through a summary article by Martinson (1974) entitled *What Works? Questions and Answers about Prison Reform*, which provided the name for the debate. In short, Martinson's answer was that nothing works. The results were reassessed through a random sample of the original literature by Fienberg and Grambsch (1979) and they found that the conclusions drawn by Lipton, Martinson, and Wilks were generally accurate. Shortly after the publication of these pessimistic conclusions, funding shifted away from rehabilitation efforts to increased law enforcement (Anstiss, 2003). This also corresponds with the time at which the incarceration rate begins to increase rapidly. In response to the claim that nothing works, several researchers were able to show that there were more positive results than negative ones among the reviewed studies and even Martinson later acknowledged that his conclusion was probably invalid (Anstiss, 2003; Maltz, 2001). Nevertheless, the Lipton, Martinson, and Wilks meta-analysis gave support to a politically conservative shift away from rehabilitation toward a more retributive-oriented approach, a swing from which the pendulum has never returned.

A second example of criminal justice policy based on recidivism rates gives a different answer to the question of what works. Murray and Cox (as cited in Maltz, 2001) studied the effect of immediate correctional intervention on recidivism among juvenile offenders. They noted a significant reduction in recidivism when there was immediate correctional intervention and concluded that getting tough works. Maltz, however, is highly critical of the study and finds that the decrease in recidivism is an artifact resulting from selection bias and behavioral peculiarities of the individuals selected for the study. Despite the efforts of other researchers to demonstrate its flaws, the study provided further support to the politically conservative approach to criminal justice policy and essentially justified the view that the solution is to lock up a ever-greater number of offenders.

In closing, a few comments should be directed toward the issue of the relevance of these examples to the current study. First, even though some of the studies mentioned by Martinson (1974) examined vocational training and found it ineffective, it appears as though there were no studies examined within the meta-analysis that focused specifically on other income-related economic variables. In general, there have been very few studies focusing on economic factors influencing recidivism.

Hence, there is considerable uncertainty regarding the effects of these types of variables on recidivism. Second, the emphasis in this study is upon policy changes that affect the economic decision making of the offenders. This type of approach is markedly different from a rehabilitation program that identifies a particular problem as a causal factor leading to criminality. As determinants of behavior, economic factors are likely among the most general determinants that can be validly applied to every member within a society. It may be that virtually everyone will have a greater tendency toward criminality when faced with material deprivation. In the final assessment, the conclusions drawn by Martinson would not appear to have significance for the current study.

CHAPTER 3

INTERNATIONAL COMPARISONS OF RECIDIVISM

A recidivism rate standing by itself cannot be evaluated with respect to its magnitude. Meaningful judgments concerning a particular recidivism rate require reference to recidivism rates drawn from other suitably comparable populations. A comparative analysis of international recidivism rates furnishes this necessary basis for evaluative purposes. The primary objective of this chapter is to determine how the recidivism rate for the United States compares with the recidivism rates of several similar countries. The assessment of the recidivism rate for the United States will be useful for developing a sense of where Utah stands in comparison to the nation when Utah's recidivism rate, as derived from the Utah parolee data, is discussed in a subsequent chapter.

A question of considerable interest, particularly with respect to the formation of criminal justice policy, is whether there are any relationships among the crime, incarceration, and recidivism rates. In addition to recidivism rates, crime and incarceration rates were collected for all of the countries selected for the international comparisons. Two relationships were examined: the relationship between recidivism and incarceration rates and the relationship between crime and incarceration rates. In each case, the relationship is found to differ from intuitive expectations and raises important questions regarding criminal justice policy in the United States.

Upon examination of the international recidivism, crime, and incarceration rates and the relationships among these rates, the United States is evaluated relative to the other countries in the comparisons. Viewing the statistics for the United

States through the lens of international comparisons, it is evident that the purpose and efficacy of criminal justice policy in the United States is in need of critical reexamination.

3.1 Recidivism in the United States and abroad

For a comparative analysis to yield useful information, the populations under consideration must be reasonably similar. In light of the fact that the recidivism rate for the United States is the central concern, the countries selected for comparison should share certain essential characteristics with the United States. Among these characteristics, those of greatest importance are the legal system, the type of government structure, the level of economic development, and the cultural background. In addition to sharing a set of requisite attributes, the countries under consideration must, of course, publish studies on recidivism. These two criteria served to determine the selection of the countries for comparison with the United States.

The natural candidates for comparison with the United States include Australia, Canada, New Zealand, and the countries of Europe. Australia, Canada, and New Zealand satisfied the two criteria and were included. However, only a relatively small number of the European countries were included because of the absence of national recidivism statistics. In a survey conducted by Wartna and Nijssen (2006), questionnaires were sent to 41 European countries, including Russia, inquiring about large-scale recidivism research. Of the 33 countries that responded to the questionnaire, only 14 reported the existence of large-scale national studies on recidivism. Even though the United Kingdom is typically viewed as a single political entity, the survey counted England and Wales, Northern Ireland, and Scotland as three individual countries. Of the 14 countries reporting the presence of national recidivism studies, only Norway was excluded due to the unavailability of crucial information regarding the nature of sentenced offenders. Thus, 13 European countries were selected for inclusion in the comparisons. The complete list of countries selected for the international comparisons along with the sources of their recidivism rates is given as follows: Australia (Jones, Hua, Donnelly, McHutchinson & Heggie, 2006; Payne, 2007), Austria (Statistics Austria, 2008), Canada (Bonta, Dauvergne & Rugge, 2003), Denmark (Denmark Department of Prisons and Probation, 2006), England/Wales (Spicer & Glicksman, 2004), Finland (Hypén, 2003), France (Kensey & Tournier, 2004, 2005), Germany (Jehle, Heinz & Sutterer, 2003; Jehle, 2005), Iceland (Baumer, Wright, Kristinsdottir & Gunnlaugsson, 2002), Ireland (O'Donnell, Baumer & Hughes, 2008), the Netherlands (Wartna, Tollenaar & Essers, 2005), New Zealand (Nadesu, 2007, 2008, 2009; New Zealand Department of Corrections, 2005), Northern Ireland (Ruddy & Brown, 2008), Scotland (Scottish Executive, 2005), Sweden (National Council for Crime Prevention, 2008), Switzerland (Storz, 1997), and the United States (Langan & Levin, 2002).

International comparisons of any type are always open to criticism on the basis that no two countries ever share enough of the necessary qualities to allow for valid comparative inferences. This view is likely correct in at least some cases. An equally strong case, however, could be made that all of the selected countries have economic, political, and cultural institutions that are quite similar to those in the United States. For the purpose of developing a general sense of where the United States stands relative to other countries in terms of recidivism, crime, and incarceration rates, the comparisons appear reasonable.

Aside from the issue of the similarity of the countries under comparison, another potentially troublesome issue is the comparability of the reported statistical figures. No single, standardized approach to the collection, analysis, and publication of recidivism statistics is used by all of the countries selected for this comparison. As a result, the differences in the statistics published by each country require a few qualifying remarks before any sound comparisons can be made. These qualifications are fully addressed below.

The recidivism statistics for the 17 countries are reported in Table 3.1. The measure of recidivism used is the reconviction rate because it is the most widely available statistic. However, only the reincarceration rates were available for Finland and Ireland. Instead of omitting these countries from the comparisons, the decision was

Table 3.1. Reconviction Rates for the Selected Countries

Country	Release Period	N	Age Range	Crime Rate[a]	Incarceration Rate[b]	Follow-up Periods in Years 1	2	3	4	5
Australia	2001-2002	2,793	18 and up	4,738	119			63.9		
Austria	2003	4,885	18 and up	4,745	96	20.7	38.4	49.4	55.5	59.8
Canada	1994-1997	14,306	18 and up	5,074	114		42.9			
Denmark	2003	7,141	15 and up	5,594	67		37.7			
Finland	1996	3,680	15 and up	4,420	66	15.1	32.1	42.1	48.2	52.2
France	1996-1997	2,408	13 and up	4,132	96					51.9
Germany	1994	22,816	14 and up	4,308	99				59.5	
Iceland	1994-1998	1,176	18 and up	2,328	38	11.0	25.0	37.0	44.0	53.0
Ireland	2001-2004	19,955	15 and up	2,578	81	27.4	39.2	45.1	49.2	
Netherlands	1996-1999	69,602	18 and up	5,758	100	43.4	55.5	61.9	66.0	69.0
New Zealand	2002-2003	4,945	18 and up	5,789	155	42.6	55.4	62.0	68.0	70.8
Sweden	2002-2005	13,958	15 and up	8,214	75			61.0		
Switzerland	1988	6,393	18 and up	4,125	71	12.3	25.5	33.8	39.7	44.5
UK: England & Wales	2001	14,569	18 and up	7,414	140		58.2			
UK: N. Ireland	2004	825	17 and up	4,684	68	21.0	44.0			
UK: Scotland	1999	5,738	16 and up	5,071	131	46.0	60.0	67.0	71.0	
USA	1994	33,796	18 and up	5,018	715	21.5	36.4	46.9		

Note. The rates reported for Finland and Ireland are reincarceration rates.
[a]Per 100,000 population, based on totals for intentional homicide, assault, rape, robbery, and theft for 2003.
[b]Per 100,000 population for 2003.

made to include their reincarceration rates and interpret them as the minima of the possible ranges of the reconviction rates. In reference to Ireland's recidivism statistics, O'Donnell, Baumer, and Hughes (2008) argue that the reincarceration rate should be close to the reconviction rate because offenders with prior convictions are more likely to receive prison sentences upon reconviction (p. 128). While arguments can be made for the inclusion or exclusion of Finland and Ireland from the comparisons, the decision was to include the statistics and leave it to the reader to choose whether or not to ignore those particular countries.

Another issue that creates some difficulty regarding comparability is the variation in the number of follow-up periods. A few countries (e.g., the Netherlands and Finland) reported statistics for as many as eight follow-up periods, while some countries (e.g., Canada and Denmark) reported a recidivism statistic for only a single follow-up period. The number of follow-up years in Table 3.1 was limited to five, even though some countries reported statistics beyond 5 years. The variation in the number of follow-up years does create some uncertainty about the exact values of the rates for those years where they are unreported. Nevertheless, there are generally enough rates reported for each of the follow-up years to allow for useful comparisons.

Of the five Nordic countries, only Norway's recidivism rate is not reported. Even though Norway has conducted a large-scale recidivism study (Statistics Norway, 2007), one crucial figure was omitted in their statistics: the number of offenders released from a custodial sentence. Thus, it was impossible to determine if individuals were released from custodial sentences or received other types of sentences. Consequently, Norway was excluded. As for the 17 countries included in Table 3.1, all of the offenders were released from custodial sentences.

One last caveat concerns the statistics for Scotland. An examination of Table 3.1 reveals that Scotland has the highest reconviction rate in each of 4 follow-up years as compared to all of the other 16 countries. However, the Scottish statistics contain a major flaw. A significant number of individuals were counted as reconvicted for offenses that were committed prior to their initial prison release. The 2-year rate is judged to be overstated by as much as 10%, but no information is provided regarding the estimated overstatement for any other year, which makes correcting the rates impossible. The rates found in Table 3.1 are the unadjusted rates taken as reported by the Scottish Executive (2005, Table 6, p. 10).

Before drawing a comparison between the United States and the other 16 countries, a few comments concerning the recidivism statistics for the United States are in order. Two large-scale recidivism studies conducted by the U.S. Department of Justice's Bureau of Justice Statistics serve as the primary references for the recidivism statistics of the United States. Beck and Shipley (1989) conducted the first study following 108,580 individuals released from prisons in 11 states in 1983. In a more recent study, Langan and Levin (2002) followed 272,111 individuals released from prisons in 15 states in 1994, which represents approximately two-thirds of all releases from prison in 1994. In both studies, the statistics were calculated using samples taken from the total number of released prisoners tracked. On the whole, the two studies are remarkably similar. In Table 3.2, three measures of recidivism are compared for the release years of 1983 and 1994. The reincarceration rates listed in Table 3.2 are based on new convictions, not technical violations. If technical violations are included, the 3-year reincarceration rate for prisoners released in 1994 is 51.8%. The reconviction rates for each year show no statistically significant differences. Because the reconviction rates are nearly the same in both studies and the 1994 study uses a larger, more recent data set, the reconviction rates reported in Table 3.1 were based on the results produced by Langan and Levin.

Turning to the comparison between the United States and the other 16 countries,

Table 3.2. United States Recidivism Rates for 1983 and 1994

	Rearrests		Reconvictions		Reincarcerations	
Follow-up Years	1983	1994	1983	1994	1983	1994
1	39.3	44.1	23.1	21.5	18.6	10.4
2	54.5	59.2	38.3	36.4	32.8	18.8
3	62.5	67.5	46.8	46.9	41.4	25.4

it must be concluded that the United States compares favorably. The United States' ranking relative to the other countries is best appreciated by directly comparing the reconviction rates in the United States with the averages for all 17 countries. The average international reconviction rates were formed by summing the rates for all of the countries that reported a rate during a particular follow-up year and dividing by the number of reported rates. This method of forming the international average was chosen over a weighted-average based on the number of observations in each country's statistics because the populations and sample sizes tended to vary widely from country to country and the desire was to give each country the same overall weight. It was previously mentioned that the recidivism measures for Finland and Ireland are reincarceration rates and that Scotland's reconviction rates included reconvictions for crimes committed prior to the most recent release. The former statistics will tend to be underestimates of the true reconviction rates for each of the two countries, while Scotland's reconviction rates are overestimated. Taken together, they will tend to offset each other. On the whole, the international averages appear to be reasonable.

Table 3.3 directly compares the average for all 17 countries with the United States for the first 3 follow-up years. The table reveals that the reconviction rate for each of the 3 follow-up years for the United States is approximately five percentage points lower than the average of the 17 countries. While the United States does not have the lowest recidivism rates in the comparisons, they are lower than those of most countries.

Two last points deserve mention before turning to the examination of the relationships among international recidivism, crime, and incarceration rates. First, it is of some interest to note how some pairs of countries that appear rather similar

Table 3.3. The International Average and U.S. Reconviction Rates

Follow-up Periods in Years	1	2	3
Average of International Reconviction Rates	26.1	42.3	51.8
United States Reconviction Rates	21.5	36.4	46.9

in terms of their economic, political, and cultural attributes have very different recidivism rates. For example, the two Scandinavian countries of Iceland and Sweden share many of the same economic, political, and cultural institutions, yet their 3-year reconviction rates are dramatically different (37% for Iceland compared to 61% for Sweden). However, some pairs of countries that share many attributes do exhibit similar recidivism rates. England/Wales and Scotland most likely exhibit a greater degree of similarity with respect to social characteristics than any other two countries in the comparisons and their reconviction rates are nearly the same: 58.2% as compared to 60%, respectively. These comments serve only to point out the precariousness of predicting a recidivism rate merely on the basis of some perception of a country's levels of economic, political, and cultural development. Second, many of the international recidivism studies included statistics related to predictors of recidivism. The consideration of these statistics is left until Chapter 4 when a more general discussion of predictors of recidivism is undertaken.

3.2 International crime, incarceration, and recidivism rates

Of the possible relationships among international crime, incarceration, and recidivism rates, two specific relationships are of primary concern. The first relationship considered is between the crime and recidivism rates and its examination is motivated by a claim that as the crime rate decreases, the recidivism rate must necessarily increase. This is of importance because the claimed relationship is based on a deterministic view of criminal behavior that precludes the ability to affect behavior through economic incentives and disincentives. The second relationship examined is between the crime and incarceration rates. The issue suggested by a consideration of this relationship is whether the percentage of the population imprisoned has any influence on the overall crime rate. Both relationships have important implications for criminal justice policy and they are examined following a few comments on the sources and problems of the statistics.

Due to the infrequent publication of recidivism statistics, it was impossible to gather recidivism rates corresponding to the same release year for all of the countries. However, crime and incarceration rates were available for the same year, the most recent of which is 2003. The incarceration rate is expressed as prisoners per 100,000 members of the population. While the incarceration rate is straightforward, the crime rate requires some explanation. The crime rate is based on the number of crimes reported to the police and is expressed as crimes per 100,000 members of the population. The crime rate is calculated using totals for intentional homicide, assault, rape, robbery, and theft. The category of theft includes burglary, motor vehicle theft, and larceny. Both the crime and incarceration rates for the European countries were taken from the *European Sourcebook of Crime and Criminal Justice Statistics, 2003* (Ministry of Justice, Research and Documentation Centre, 2006). For Australia and Canada, the source of the crime and incarceration rates was the *Ninth United Nations Survey of Crime Trends and Operations of Criminal Justice Systems, Covering the Period 2003-2004* (United Nations, 2007). New Zealand's crime rate was taken from a report published by the New Zealand Police (2004) and the incarceration rate was taken from the *Census of Prison Inmates and Home Detainees, 2003* (New Zealand Department of Corrections, 2004). For the United States, the incarceration rate is reported in Harrison and Karberg (2004) and most of the crime rate figures were taken from *Crime in the United States, 2003: Uniform Crime Reports* (Federal Bureau of Investigation, 2004). Only one statistic for the United States' crime rate requires additional explanation. The uniform crime reports do not report total assaults, but only aggravated assaults. An estimate of total assaults was based on figures drawn from the *National Crime Victimization Survey* (Catalano, 2004). The estimated crime rates found in the *National Crime Victimization Survey* tend to be higher than those based on police report data because not all victims report the crimes committed against them. Therefore, the figure for total assaults found in Catalano was reduced by roughly 28% to reflect similar differences among the crime rates found in the police report data and victimization self-report data for other types of crimes.

Interest in the investigation of the relationship between crime and recidivism rates is stimulated by a claim that there is a negative relationship between these two variables. Kanazawa (2008) states that a trade-off exists between the crime rate and the recidivism rate in the following remarks:

> So regardless of how tough the law enforcement or how effective the prison system, the lower the crime rates, the higher the recidivism rates in any society at any time. You can have one or the other, but not both at the same time. (¶ 7)

Kanazawa's view is based on an evolutionary psychology interpretation of a theory put forth by Moffitt (1993) that asserts that all crimes are committed by only two types of individuals. The first type of criminal engages in adolescence-limited antisocial behavior, which is to say that criminal activity for these individuals occurs only during late adolescence and early adulthood and ceases thereafter. The second type engages in life-course-persistent antisocial behavior, which implies that these individuals will commit crimes throughout their entire lives. Whereas the volume of crime committed by the adolescence-limited types can fluctuate from year to year depending on various social factors that might influence adolescents and young adults to commit crime, the number of life-course-persistent types in a society and the volume of crimes they commit are virtually constant over time. Therefore, a reduction in the overall crime rate reflects a decrease in the number of adolescence-limited types committing crimes leaving a higher proportion of life-course-persistent types in the total population of criminals, thereby increasing the recidivism rate.

This assertion is important for criminal justice policy because it makes a rather strong claim about the possibilities of behavioral change of criminal offenders. While Kanazawa allows for the possibility that the overall level of crime resulting from adolescence-limited antisocial behavior can fluctuate, he asserts that the number of those exhibiting life-course-persistent antisocial behavior is genetically determined and, therefore, immutable. If this view is correct, it follows that recidivism cannot be influenced in any way through modifications of criminal justice policy. Moreover, it denies not only the possibility that an offender could have chosen not to recidivate, but that offenders can modify their behavior in response to economic incentives and disincentives. If empirical evidence contradicts Kanazawa's claim, then some measure of doubt must be cast upon his deterministic view and the relative merit of the choice-theoretic perspective must consequently rise in estimation.

The examination of the relationship between international recidivism and crime rates is performed at an elementary level and is purely descriptive in nature. Table 3.1 indicates that 13 countries reported a 2-year reconviction rate and these countries were used for examining the relationship between recidivism and crime. A scatterplot for these 13 countries is presented in Figure 3.1. All of the crime rate statistics are dated to 2003, but the dates for the 2-year reconviction rates vary from country to country. Because the prisoner release dates for the reconviction statistics tend to be relatively close to 2003 and reconviction rates do not appear to fluctuate radically from year to year, the reconviction rates used to examine the relationship appear reasonably accurate. Figure 3.1 does not produce any visual evidence that the relationship between recidivism and crime is negative.

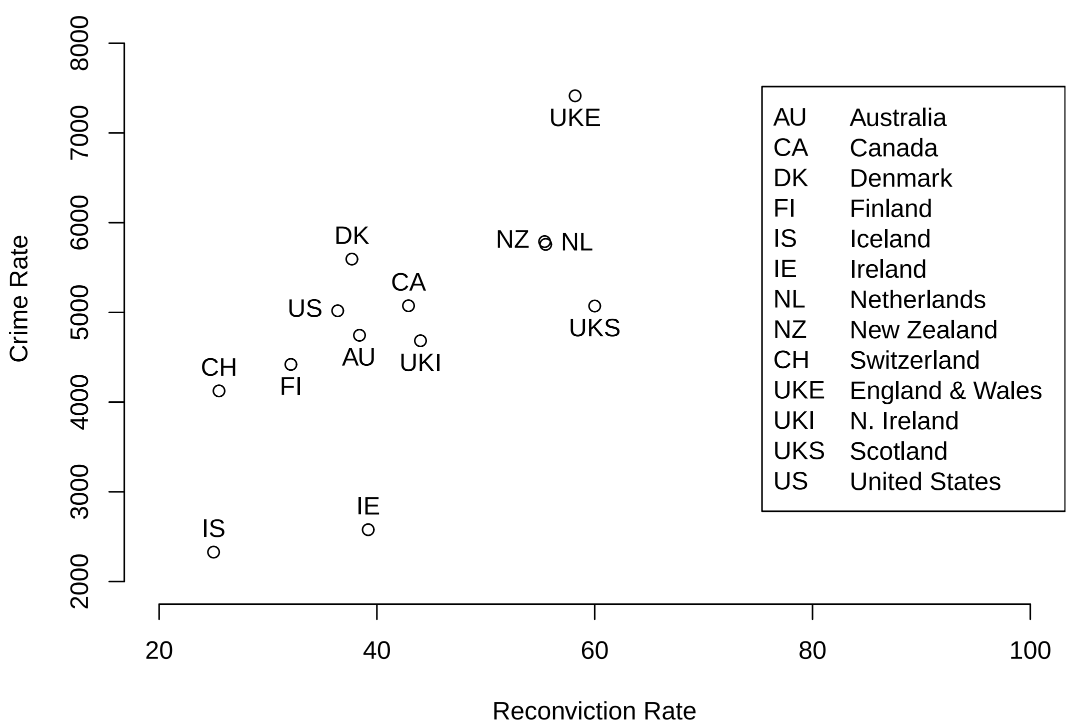

Figure 3.1. Scatterplot of International Reconviction and Crime Rates

More formally, the Pearson correlation coefficient for this relationship is .6992 with $p = .007$, which indicates a rather strong positive relationship. Thus, the evidence appears to contradict Kanazawa's claim that the relationship between recidivism and crime is negative and thereby undermines this particular deterministic theory of recidivism behavior.

The second relationship of interest concerns the crime and incarceration rates. This relationship is of great importance to criminal justice and corrections policy because its true direction and strength can either support or refute the proposition that incarcerating a greater number of offenders lowers the overall crime rate. As with the analysis of the previous relationship, the examination of the relationship between crime and incarceration is elementary and purely descriptive. The statistics for both the crime and incarceration rates are all dated to 2003 for all countries. All 17 countries are represented in the scatterplot in Figure 3.2. A fitted line was included to provide a sense of the direction of the numerical relationship.

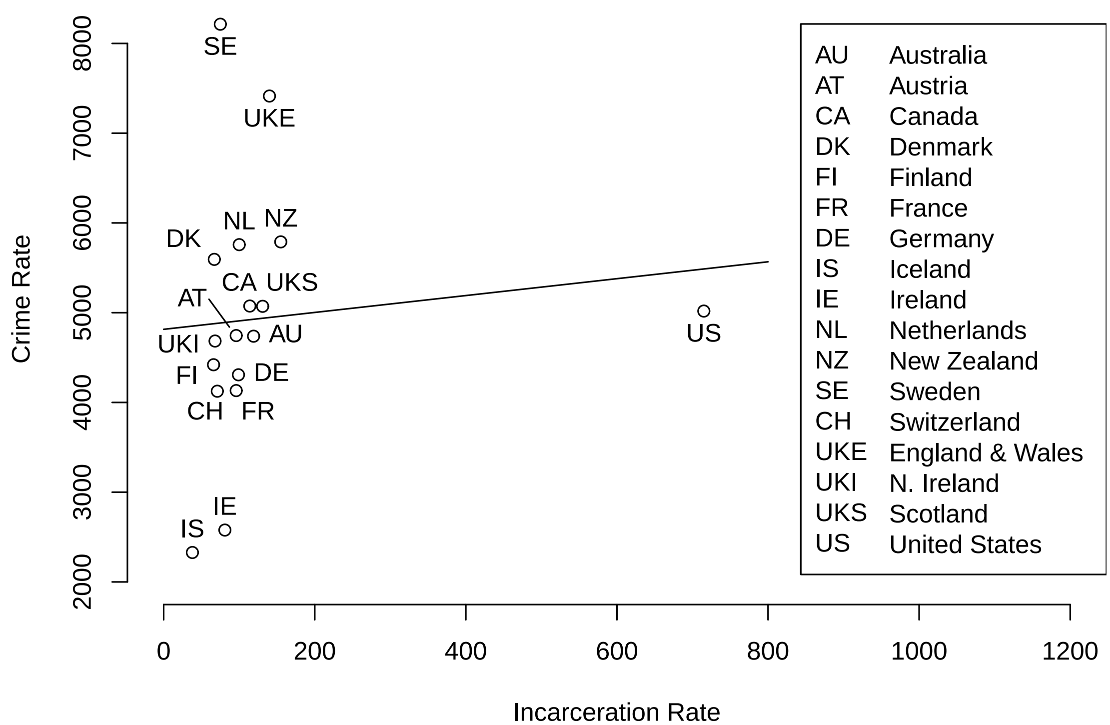

Figure 3.2. Scatterplot of International Crime and Incarceration Rates

As Figure 3.2 makes apparent, one particular country stands out from the rest. The scatterplot reveals that, in comparison with all of the other countries, the United States has an extremely high incarceration rate. When the United States is excluded from the calculation, the mean incarceration rate per 100,000 members of the population for the other 16 countries is 94.75. With an incarceration rate of 715 per 100,000 members of the population, the incarceration rate for the United States is more than seven times the size of the mean incarceration rate for the other 16 countries.

The United States is such an extreme observation that the other observations are consequently compressed in the graph, thereby making it difficult to ascertain the nature of the relationship. The fitted line reveals that the relationship is slightly positive. The correlation coefficient is .0997 and the probability is very low that a relationship exists at all. If the United States is excluded, the correlation coefficient is .4367 and it is likely that the relationship is, in fact, positive.

This elementary analysis suggests that the policy of attempting to lower the crime rate by imprisoning an ever larger numbers of individuals has no empirical support. The mean crime rate for all 17 countries is 4,941 per 100,000 members of the population and the United States is very close to the mean with a crime rate of 5,018 per 100,000. With such a high incarceration rate, if there were any substance to the claim that locking up more prisoners reduces the crime rate, the United States should exhibit a crime rate substantially below the mean. It would be presumptuous to assert that those who determine criminal justice and corrections policy base their policy decisions on statistical evidence, regardless of what prior statistical evidence might suggest about this particular relationship. Nevertheless, the high incarceration rate in the United States is likely to find its justification in either a belief that incarcerating more individuals reduces crime or a belief that retribution demands a high level of incarceration. If the rationale behind locking up large numbers of individuals in prison is the belief that it lowers crime, the evidence does not appear to support this view. If retribution is the motive behind the high incarceration rate, the satisfaction of this subjective moral belief carries a rather high price in terms of taxpayer dollars.

While the analysis of the relationship between the crime and incarceration rates found above does not constitute a definitive proof of the relationship's qualitative or quantitative properties, it certainly raises questions about justification and efficacy of policies that support increasing the incarceration rate. Clearly, criminal justice and corrections policy must address two important issues. First, the relationship between the crime and incarceration rates must be extensively researched to definitively establish its direction and magnitude so as to determine its role within the formation of policy decisions. Second, a rational, well-informed debate is needed to determine whether the public interest is best served by policies motivated by retribution or policies focusing on public safety and cost minimization.

3.3 Evaluating the international standing of the United States

The absence of standardized, internationally-accepted methods for collecting and analyzing recidivism data makes international comparisons precarious at best. For two countries, Finland and Ireland, only reincarceration rates were available and these rates had to be interpreted as lower bounds for the reconviction rates. Scotland's data contains errors in the form of "pseudo-recidivism," reconvictions for crimes committed prior to the initial prison sentences from which the offenders were released. While the three aforementioned countries put forth rates that were approximations of reconviction rates, some doubt remains as to their overall accuracy. With the recent appearance of higher quality, nationwide recidivism studies, it appears as though many countries in Europe are moving toward adopting the policy goal of producing annual, standardized recidivism statistics. At the present, there are still some problems of comparability with respect to international recidivism rates. These difficulties notwithstanding, the international comparisons of recidivism are viewed here to be reasonably accurate and useful for evaluating how the United States ranks in comparison with other countries exhibiting similar economic, political, and cultural qualities.

In summary, the recidivism rate in the United States, measured in terms of the reconviction rate, was found to be slightly lower than the average of the 17 countries in these comparisons. Figure 3.2 is an accurate graphical representation of where the United States stands in relation to the highest and lowest reconviction rates for the first 3 follow-up years. Because the recidivism statistics for the Netherlands are of the highest quality, their reconviction rate was used to represent the highest rate. While Iceland had a slightly lower reconviction rate during the first 2 follow-up years, Switzerland's reconviction rate was chosen to represent the lowest overall rate because its 5-year rate was clearly the lowest for all 17 countries. In evaluative terms, the United States' recidivism rate compares favorably with most of the selected countries. No attempts have been made within the literature to determine the lowest recidivism rate that a given society could realistically achieve. However, the United States' 3-year reconviction rate is 38.8% larger than that of Switzerland, which suggests that there is room for improvement for the United States.

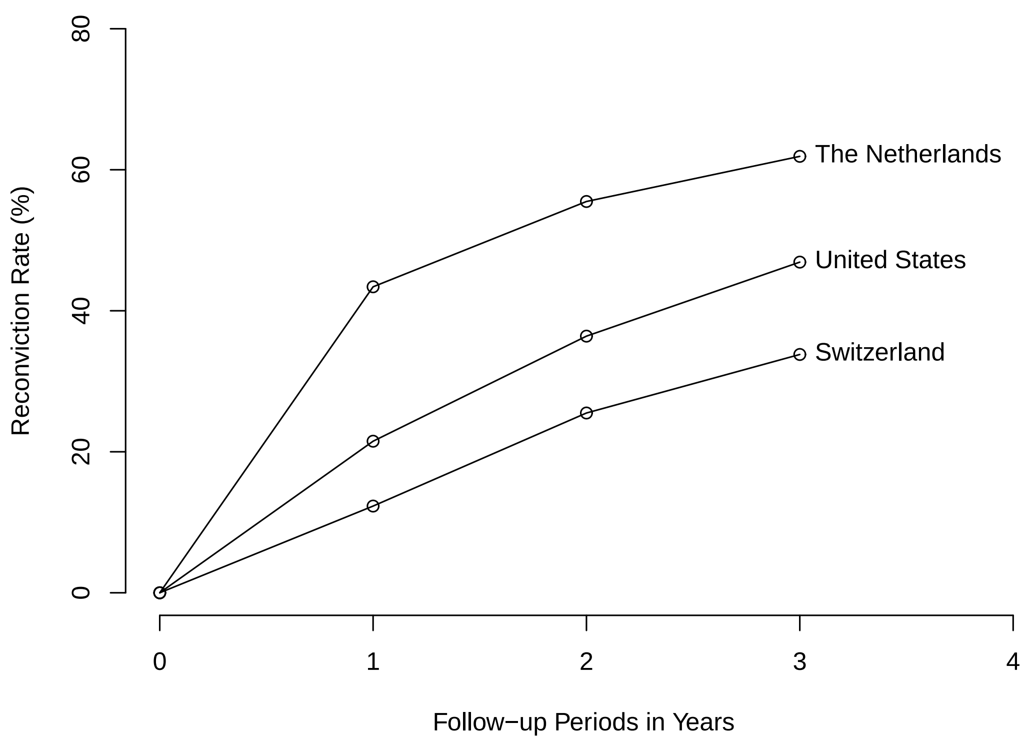

Figure 3.3. Reconviction Rates for The Netherlands, Switzerland, and the U.S.

Besides the international comparisons of recidivism, two relationships between crime, incarceration, and recidivism rates were examined in this chapter. The first relationship examined was between the crime and recidivism rates and the empirical evidence appears to contradict a deterministic theory that asserts a negative relationship. Showing that this theory may not have any empirical support is important because it might ultimately change some opinions toward the view that successful rehabilitation and reintegration is still a possibility. The second relationship considered was between the crime and incarceration rates. The belief that the crime rate decreases as more criminals are incarcerated implies a negative relationship between these variables. The evidence indicates that either there is no relationship between these variables or the relationship is, in fact, positive. In either case, the evidence here does not support the view that increasing the incarceration rate lowers the crime rate.

CHAPTER 4

PREDICTORS OF RECIDIVISM

An examination of predictors of recidivism from past studies is conducted in this chapter. With the vast number of recidivism studies that have been carried out in the past, no pretense is made for this being an exhaustive survey. The sources include academic meta-analyses, academic journal articles, and reports issued by government justice and corrections departments. Statistical information from the recidivism studies of the 16 foreign countries that were included in the international comparisons is also used. The results taken from the international studies are provided in summary form without direct citation. The sources for all of the statistics taken from the studies used in the international comparisons are found in Chapter 3.

To emphasize the difference between past recidivism studies and the focus of this dissertation, the predictors are grouped into two categories: economic and sociological predictors. The term *economic predictor* refers to a factor that is more or less directly related to income. These factors include employment status, wage rate, job type, restitution payments, child support payments, and so forth. Educational predictors are probably best classified as economic predictors because education can be viewed as the accumulation of human capital, which has a direct relationship to income. The most important characteristic of an economic factor is that it conveys information regarding incentives and disincentives toward particular types of behavior. In contrast, the term *sociological predictor* refers to a factor that is not directly related to income. Sociological predictors include such factors as age, gender, and race, among others. While sociologists have attempted to explain behavior in terms of these types of variables, such factors play a much smaller role in the analysis of behavior from the economic point of view. Besides the explanatory differences associated with each type of predictor, there are two practical reasons for drawing the distinction between sociological and economic predictors. First, sociological information is relatively easy to collect from corrections departments and, therefore, is relatively inexpensive to collect. Economic information, on the other hand, is more expensive to gather. In order to justify the additional cost associated with collecting economic information, there must be reason to believe that economic factors are important in predicting recidivism. The second reason for separating economic and sociological predictors concerns their relative usefulness for policy purposes. Because economic factors are changeable, criminal justice policy can target these variables in order to influence behavior. Sociological variables, on the other hand, are either unchangeable accidents of birth or non-behavioral historical facts, neither of which can be modified through policy.

While many factors have been examined in the past with respect to their relationship to recidivism, the focus here centers on those factors for which information was collected from parolees in Utah. Each of the first two sections begins with a list of the statistical information collected from the Utah parolees and is followed by a survey of past results for each factor. The last section of this chapter briefly summarizes the findings.

4.1 Sociological predictors of recidivism

Information was collected on age, gender, race, prior incarcerations, and most recent crime committed prior to release from the Utah parolees. Only these five predictors are surveyed in this section. Age, gender, and race information is sociological in the sense that there might be some relationship to recidivism that can be explained by way of a sociological theory, but a direct relationship to economic theory is less clear. Prior incarcerations and the type of most recently committed crime are also sociological in nature because they express historical facts that do not always have a clear relationship to economic incentives and disincentives. Moreover, they appear irrelevant to forward-looking, optimizing behavior.

Regarding the age at time of release from prison, many studies indicate that there is a strong negative relationship between age and recidivism (Bales & Mears, 2008; Beck & Shipley, 1989; Chiricos, Barrick, Bales & Bontrager, 2007; Gendreau, Little & Goggin, 1996; Langan & Levin, 2002). Pritchard (1979) notes that age at first arrest has been a strong predictor in the majority of past recidivism studies, which is further substantiated by Beck and Shipley. From the international recidivism studies, 14 of 16 countries provided recidivism rates according to age and in 13 countries the recidivism rates were found to be highest for those between the ages of 15 and 30, exhibiting a steady decline thereafter. The relationship between age and recidivism was slightly different for Sweden, where the decline in the recidivism rate did not occur until after 50 years of age. The past studies provide convincing evidence for a negative relationship between age and recidivism.

Turning to the relationship between gender and recidivism, a large number of studies have shown that males are more likely to recidivate as compared to females (Bales & Mears, 2008; Beck & Shipley, 1989; Chiricos et al., 2007; Gendreau et al., 1996; Langan & Levin, 2002). Statistical information on gender and recidivism was reported in 13 of the 16 international studies and in 10 countries females had a clearly lower recidivism rate as compared with males. However, in three countries, the recidivism rate was either not statistically different between males and females (Australia) or females exhibited a higher rate of recidivism (Northern Ireland and Sweden). Thus, the relationship between gender and recidivism is somewhat strong, but not always statistically significant.

The summarizing of race variables is complicated in part by the distinctions between race and ethnicity. The term *race* is usually applied to the distinction between Blacks and Whites, while *ethnicity* typically involves classifying individuals as either Hispanic or non-Hispanic. Thus, race and ethnicity are neither interchangeable nor mutually exclusive. Several studies indicated that Blacks are more likely to recidivate compared to Whites (Bales & Mears, 2008; Beck & Shipley, 1989; Chiricos et al., 2007; Langan & Levin, 2002). Beck and Shipley found that Hispanics are more likely to recidivate than non-Hispanics, but two later studies showed that Hispanics are less likely to recidivate compared to non-Hispanics (Chiricos et al.; Langan & Levin). In the meta-analysis by Gendreau et al. (1996), race variables were generally found to be significant predictors of recidivism, while in the meta-analysis by Pritchard (1979), race was considered insignificant because half the studies showed a relationship and half did not. Only Australia, Canada, and New Zealand from the 16 international studies reported recidivism statistics by race and in all three cases they showed that the indigenous peoples had higher recidivism rates than those of European ancestry. The past evidence appears to indicate that only particular races (e.g., Blacks and indigenous peoples) exhibit a relationship to recidivism. When other races are grouped together (e.g., Whites and Hispanics), no difference in recidivism appears to exist.

The number of prior incarcerations is a universally strong predictor of recidivism (Bales & Mears, 2008; Beck & Shipley, 1989; Chiricos et al., 2007; Gendreau et al., 1996; Langan & Levin, 2002; Pritchard, 1979). In all of these past studies, there is a positive relationship between the probability of recidivism and prior incarcerations. In 12 of the 16 international recidivism studies, recidivism rates were provided for released prisoners according to the number of prior incarcerations. In every study, there was a strong positive relationship between these two variables.

The type of crime that was most recently committed before release is strongly related to recidivism and varies according to type of crime. In the meta-analyses of Gendreau et al. (1996) and Pritchard (1979), crime type is found to be generally significant. Pritchard notes that those convicted of motor vehicle theft have a higher recidivism rate than those convicted for other crimes. Released prisoners convicted of a property crime, such as burglary, theft, motor vehicle theft, or dealing in stolen property, exhibit higher recidivism rates relative to other crimes (Beck & Shipley, 1989; Chiricos et al., 2007; Langan & Levin, 2002). In 11 of the 16 international recidivism studies, those convicted of theft, dealing in stolen property, and burglary all have higher recidivism rates as compared to those convicted of other crimes. The crime types associated with the lowest recidivism rates are homicide and sexual offenses (Beck & Shipley; Langan & Levin). In the 11 international recidivism studies that provided information about recidivism with respect to crime type, homicide and sexual offenses consistently exhibited the lowest recidivism rates.

4.2 Economic predictors of recidivism

The economic information collected from Utah parolees included employment status, wage rate, hours worked per week, job type, employer-provided health benefits, presence of a second job, restitution payments, and child support payments. Data was also collected on living conditions of parolees, which included rent, number of roommates, and assistance received for paying rent. While some results from previous studies exist for a few of these factors, the number of studies is small. In the case of some economic factors, no previous studies could be found that studied the variables in question. Consequently, the following survey of past results on economic predictors of recidivism is rather sparse as compared to the survey of sociological predictors.

With respect to employment, Pritchard (1979) found that employment stability was a very significant predictor of recidivism, while employment status was not. Pritchard's conclusion was based on 60 studies conducted before 1979, where only 40 indicated significant results. It would seem that Pritchard could be accused of being overly pessimistic in his conclusion considering that 40 significant results out of 60 could be interpreted as evidence that there is some, albeit weak, relationship to recidivism. However, in two studies of drug offenders, Sung (2001) found that employment significantly reduced recidivism, while Kim, Benson, Rasmussen, and Zuehlke (1993) found no relationship between employment and recidivism. Only one of the international studies contained information regarding recidivism and employment. In the French recidivism study, Kensey and Tournier (2004, 2005) found that those who were employed had a much lower rate of recidivism compared to the unemployed. From this limited set of past studies, the results do not present strong evidence that employment is related to recidivism.

Turning to other employment-related variables, Pritchard (1979) reports that 11 of 15 studies have shown that wages are related to recidivism and that 13 of 19 studies have shown that job type is related to recidivism. Although the number of studies upon which these results are based is rather small, there seems to be an indication that wages and job type have some relationship to recidivism. Pritchard, however, does not specify the types of jobs considered and how recidivism rates vary according to job type. In a study that examines only drug offenders, earnings, which will be closely related to wages, were not found to be significant (Kim et al., 1993). Pritchard also notes that living arrangements have some impact on recidivism, but there is no indication of what factors were actually measured.

While there have been several studies that have examined restitution and its effect on recidivism, it is difficult to draw any general conclusions. Heinz, Galaway, and Hudson (1976) claim that recidivism is lower for those who pay restitution, but there are several reasons for skepticism regarding their claim. With a total of only 36 observed individuals, the sample size in their study is very small. Furthermore, those paying restitution were selected from a special program where individuals resided at a restitution center, a situation that can produce biased results. Ruback and Bergstrom (2006) found that those who paid restitution had lower rearrest rates than those not paying restitution. In a four-county study, Schneider (1986) found that juveniles who paid restitution had fewer contacts with courts as compared to those who served detention or were placed on probation, but the results were significant for only two of four counties. In a meta-analysis conducted by Latimer, Dowden, and Muise (2005), they found that those who paid their restitution were less likely to recidivate. However, Latimer et al. note that in all studies examined, the individuals volunteered to participate in a restorative justice program, which raises the issue of a self-selection bias. The validity of the results from these past studies of restitution and recidivism is certainly open to criticism. Nevertheless, taking these results at face value, the studies indicate a somewhat weak negative relationship between restitution and recidivism.

Concerning educational variables, Beck and Shipley (1989) found that there is a significant negative relationship between educational attainment and recidivism. Pritchard (1979), however, notes that a greater number of past studies show that education is not related to recidivism as compared to those that show a relationship. Zgoba, Haugebrook, and Jenkins (2008) report that GEDs acquired in prison do not have any relationship to recidivism. This limited sample of results suggests that there is no strong relationship between education and recidivism.

Studies could not be found that examined hours worked per week, second jobs, employer-provided health benefits, or child support payments and their impacts on recidivism. Thus, the relationships between these variables and recidivism are unknown.

4.3 A summary of the best predictors

The importance of the predictors are summarized according to whether the past studies showed a strong, weak, insignificant, or unknown relationship to recidivism. Among sociological variables, past studies provide consistent evidence that age, race, prior incarcerations, and most recent crime committed before release exhibit a strong relationship to recidivism. Gender and Hispanic ethnicity have shown mixed results in relation to recidivism.

None of the past studies surveyed here have shown strong evidence of relationships between economic predictors and recidivism. Studies have shown that wages, job type, and living arrangements are related to recidivism, but the number of past studies is small. Several studies have indicated that restitution payments reduce recidivism, but there are some concerns regarding the validity of these results due to small sample size and questionable sampling methods. Employment status and educational attainment have shown mixed results, so there is some uncertainty as to their importance to recidivism. However, there appears to be no relationship between GEDs acquired in prison and recidivism.

No studies were found that examined hours worked per week, second jobs, employer-provided health benefits, or child support. Regarding these variables, their relationships to recidivism are unknown.

CHAPTER 5

THE ANALYSIS OF THE UTAH PAROLEE DATA

The statistical analysis of recidivism among parolees released from state prison in Utah is presented in this chapter. The first section describes the survey instrument, the sampling methodology, the data cleaning process, the choice of predictor variables, and the descriptive statistics of the parolees. A brief discussion of the statistical modeling techniques used is found in the second section. The results of the statistical analyses are contained in sections 5.3 through 5.5. The statistical analysis focuses upon the use of two Bayesian techniques: Bayesian Model Averaging and Classification and Regression Trees. These techniques are were chosen to address the issues of model specification uncertainty and nonhomogeneous relationships in the data, respectively. The analysis centers upon identifying which variables belong to the best model of recidivism and comparing the effectiveness of economic and noneconomic variables as predictors of recidivism. Section 5.6 contains the generalizations of the analysis.

5.1 The survey and descriptive statistics

The Utah 2006 Census of Parolees served as the primary source of the data used in this analysis. Upon testing the survey instrument, the survey was administered to all of the Utah parolees that reported to their parole officers during the third week of May, 2006. All parolees in Utah are required to report on a monthly basis, which implies that approximately one-fourth of Utah's parolees were surveyed. The assignment of reporting dates to parolees within any month appears to be completely arbitrary, suggesting a high likelihood that the sample is perfectly random. The response rate to the survey was 100%. Given the large size of the sample relative to the population, its apparent randomness, and the 100% response rate, it can be confidently asserted that the sample is an unbiased, representative sample of the Utah parolee population.

Two versions of a questionnaire were used, one for the employed and one for the unemployed. The information gathered falls into three broad categories: demographic, educational, and economic. The demographic information collected included age and race. The educational information gathered included educational attainment (GED, high school diploma, college degree, and vocational certificate) and participation in prison educational programs (GED, high school, postsecondary, or none). The economic information can be separated into two categories: work-related information and information about the parolee's living conditions. For employed parolees, data was collected on their hourly wage, hours worked per week, occupation, receipt of health benefits, and employment at a second job. Unemployed parolees were asked about reasons for not having a job, time spent looking for employment, methods used to find employment, and type of work sought. With respect to living conditions, information was gathered from all parolees regarding rent, number of people living in the household, and the receipt of assistance for rental payments. All parolees were also asked if they owned a car. Finally, information was collected for all parolees on restitution payments, child support payments, and the amounts of the payments in each case.

The information gathered from the survey was combined with information provided by the Utah Department of Corrections. The data from the Utah Department of Corrections included information on gender, number of prior incarcerations, type of crime for which the parolee was most recently incarcerated, and returns to prison after May 2006. These two sources provided all of the data used in the analysis below.

Data was collected from 760 parolees. However, not all of the initial 760 observations were usable for the statistical analyses. Three-year follow-up information was only available for 666 of the initial 760 observations. This was due to the parolee either not indicating his or her Utah State Prison (USP) number or incorrectly recording the number. Incorrect USP numbers were detected when matching the survey data to the corrections data and these individuals were eliminated. Another 18 observations were eliminated because the individual in each case either died and was not reimprisoned, was never previously imprisoned, or was imprisoned during the period at which the survey was administered. In the latter two cases, these are likely the results of parolees incorrectly recording their USP numbers. This further reduced the number of usable observations to 648.

Additional observations were excluded due to missing values in the data set. At this point in the process, personal judgment enters as a choice must be made between the trade-off of either removing observations or removing variables from the data set. Some variables were considered crucial to the analysis, while others had to be eliminated for the sake of maintaining the number of observations. Very few parolees had second jobs, so all of the variables related to having a second job were excluded. Likewise, all of the living conditions variables were left out because there were many missing values. Finally, the variables recording the amounts for restitution and child support payments were also eliminated due to the large number of missing values. However, the dummy variables for restitution and child support variables were retained. After eliminating the variables mentioned above and excluding all observations with missing values, 506 observations remained for use in the data analyses.

The descriptive statistics for the 506 observations are produced in Table 5.1. A few of the variables are in need of some additional explanation as their use in the modeling process required special coding. For the wage and the hours worked per week variables, these were coded as 0 for those who were unemployed. As a result, the mean values for these variables in Table 5.1 are lower than the mean values for only those who were employed. For the 383 parolees who were employed, their mean wage was $10.67 and their mean number of hours worked per week was 39.4. The job type variables also require some clarification. There are eight job types, all of which are dummy variables, but their sum equals 86% of the data. This is higher than the percentage of those employed (76%) because it was possible that

Table 5.1. Descriptive Statistics for Utah Parolees

Variable	Minimum	Maximum	Mean	Median
Age at time of survey	20.58	74.92	36.87	35.66
Male	.00	1.00	.86	1.00
African American	.00	1.00	.06	.00
Asian	.00	1.00	.01	.00
Hispanic	.00	1.00	.15	.00
Native American	.00	1.00	.02	.00
Pacific Islander	.00	1.00	.02	.00
White	.00	1.00	.73	1.00
Prior Incarcerations	1.00	9.00	1.98	1.00
Most recent crime: Driving	.00	1.00	.06	.00
Most recent crime: Drug offense	.00	1.00	.25	.00
Most recent crime: Murder	.00	1.00	.03	.00
Most recent crime: Other	.00	1.00	.02	.00
Most recent crime: Person	.00	1.00	.17	.00
Most recent crime: Property	.00	1.00	.27	.00
Most recent crime: Sex offense	.00	1.00	.20	.00
Most recent crime: Weapons	.00	1.00	.01	.00
GED	.00	1.00	.36	.00
High school diploma	.00	1.00	.71	1.00
College degree	.00	1.00	.09	.00
Vocational certificate	.00	1.00	.21	.00
Prison education: None	.00	1.00	.27	.00
Prison education: GED	.00	1.00	.23	.00
Prison education: High school	.00	1.00	.50	1.00
Prison education: Post-secondary	.00	1.00	.30	.00
Employed	.00	1.00	.76	1.00
Hours per week	.00	84.00	29.85	40.00
Wage	.00	47.00	8.08	8.50
Health benefits	.00	1.00	.26	.00
Job type: Management	.00	1.00	.05	.00
Job type: Building	.00	1.00	.04	.00
Job type: Sales	.00	1.00	.07	.00
Job type: Office	.00	1.00	.11	.00
Job type: Construction	.00	1.00	.30	.00
Job type: Installation	.00	1.00	.10	.00
Job type: Production	.00	1.00	.12	.00
Job type: Transportation	.00	1.00	.07	.00
Restitution payments	.00	1.00	.50	1.00
Child support payments	.00	1.00	.33	.00
Car ownership	.00	1.00	.42	.00

some parolees held more than one job.

The descriptive statistics and their interpretations for the remaining variables are straightforward. The sociological variables are discussed first. Of the 506 parolees in the final data set, the average age was 36 and 86% were male. The race dummy variables exclude those who did not respond, so the sum will equal 100% of the data. Whites (73%) and Hispanics (15%) were the largest groups. Every parolee had at least one prior conviction and the average was two prior convictions. Regarding the last crime for which the individual was incarcerated, the two most common crimes were property crimes (27%) and drug-related crimes (25%). The least frequent crimes were weapons-related (1%), other crimes (2%), and murder (3%).

Turning to the educational variables, 71% of the parolees had a high school diploma and 9% had a college degree (the latter variable includes 2-year associates degrees). Approximately 50% attended high school courses while in prison, 30% attended postsecondary courses, and 27% did not participate in prison educational programs at all. It should be noted that both the educational attainment and prison education categories are not mutually exclusive. Thus, the variables GED, high school diploma, college degree, and vocational certificate sum to greater than 100%. The same is true for the prison education variables GED, high school, post-secondary, and none.

The only economic variables not mentioned thus far are health benefits, restitution payments, and child support payments. The health benefits variable specifically measures employer-provided health care benefits. For those who were employed, 34.5% received benefits. Roughly half of the 506 parolees made restitution payments and one-third made child support payments.

Finally, a few words should be devoted to characterizing the recidivism variable. Recidivism was measured as any return to prison, whether a new crime or technical violation, during the 3-year period after the survey was administered. Of the 506 parolees used in the final data set, 221 returned at least once to prison. This yields a 3-year reincarceration rate of 43.8% for Utah. For the purpose of comparison, the 3-year reincarceration rate, based on both new crimes and technical violations, is 51.8% for the United States (Langan & Levin, 2002). Figure 5.1 shows the percentage of parolees reincarcerated over the 3-year period. The shape of the graph is very typical as compared to those from other recidivism studies. It is a well-established fact that the largest number of offender who do return to prison will do so during the first year, with a smaller percentage returning each subsequent year. Using more precise terminology, the recidivism rate can be describe as increasing as a decreasing rate.

5.2 The statistical methodology

For the prediction of recidivism, the relevant type of model is one designed to predict the value of a binary dependent variable. The basic set of standard models used to predict binary dependent variables includes the linear probability model, the logit model, and the probit model. While one of the standard techniques is used here to model the data, the central focus rests upon two Bayesian modeling

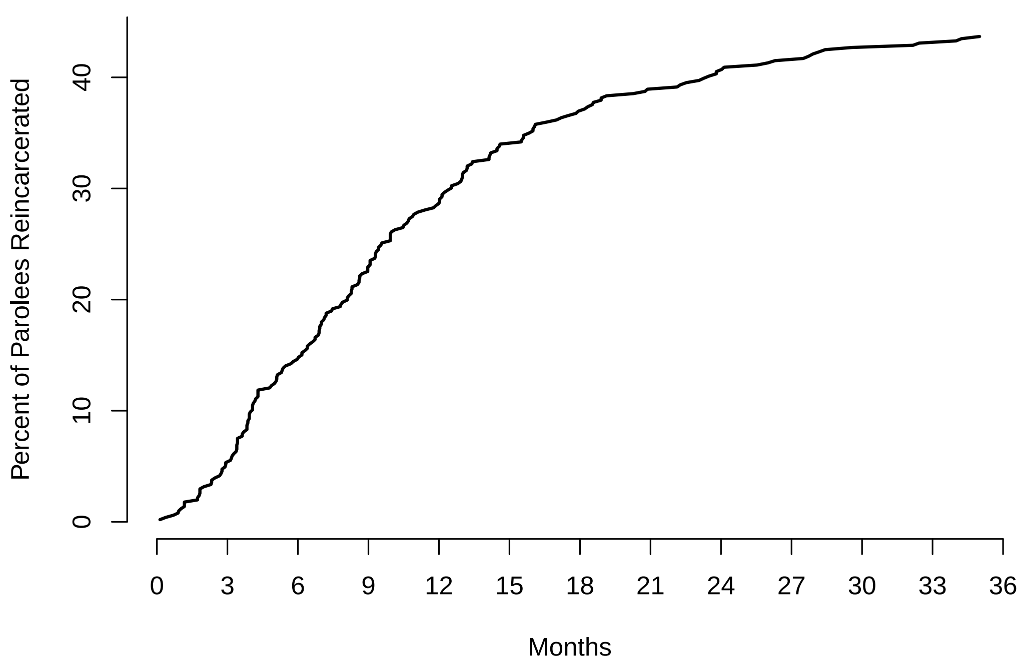

Figure 5.1. Utah Parolee Recidivism Over Time

techniques. The only standard model used to analyze the Utah parolee data is a linear probability model and it is included strictly for the purpose of comparison. Besides the traditional criticisms of the linear probability model, which include predicted probability values great than one or less than zero and inherent heteroskedasticity, there are more serious difficulties that cannot be solved by this modeling technique. In and of itself, a classical linear regression (CLR) model estimated using ordinary least squares (OLS) offers no guidance regarding the issue of model specification uncertainty. Once the analyst chooses a set of independent variables, an OLS regression estimates the coefficients as if the specification is, in fact, the best model. Although there are tests designed to detect specification errors in CLR OLS models (e.g., omitted variable and RESET tests), they are designed to signal that the model is misspecified without indicating where the error lies. One of the Bayesian methods used in the analysis is specifically directed toward this problem.

A second difficulty involves modeling nonhomogeneous relationships within the data. An OLS regression estimates coefficients using every observation in the data set, which forces it to ignore differences in a relationship across the range of the data. It is true that linear regression models can deal with misspecified functional form through creative data transformations, but, for data sets with large numbers of predictors, potentially hundreds of new variables may result. This, in turn, only exacerbates the model selection problem. The other Bayesian technique used here specifically addresses the problem of nonhomogeneous relationships in the data.

The statistical analysis of the data centers upon two Bayesian techniques. The first technique is Bayesian Model Averaging (BMA) and its importance is in addressing the two problems of model specification uncertainty and coefficient magnitude uncertainty. This technique is fundamentally important to the analysis of recidivism because there is no consensus on the specification of the best recidivism model. Tests such as the omitted variable and RESET tests give only a general indication that a model might be misspecified and they point to unknown variables outside of the model. The BMA approach to model specification uncertainty is quite different. BMA calculates posterior distributions for each of the coefficients so that a probability can be assigned to each variable indicating the likelihood that it is a predictor in the best model. It is in this way that it addresses model specification uncertainty. The method also takes into consideration uncertainty regarding the size of the coefficient. If a variable is included in a model and it is less than certain that it belongs to the best model, the estimated coefficient should reflect that uncertainty. BMA addresses this issue by adjusting the coefficients in accordance with their probabilities of being members of the best model.

The second Bayesian technique used here is Classification and Regression Tree (CART) analysis. The tree models produced by this method are easy to interpret and use in comparison to standard regression output. More importantly, the CART method can produce tree models that fit the data better and offer more accurate predictions by exploiting nonhomogeneous relationships in the data. An OLS regression operates on only the entire range of values for the variables and does not take into account changes in the relationships across the ranges of the data. A tree model is essentially a set of conditional statements, which make this a fundamentally Bayesian approach. The conditional statements that determine the branching of a tree model effectively separate the data into subsets. Each subset may reveal new relationships among the variables that were not discernable at the level of the entire data set. Subsequent branches are formed by selecting new conditional statements based on these new relationships that separate the data in a way that improves prediction. Thus, the CART method effectively addresses the issue of nonhomogeneous relationships, an issue that presents significant difficulties for OLS regression models.

5.3 Bayesian Model Averaging

5.3.1 Introduction and theoretical overview

To motivate the problem of model specification uncertainty, note that the Utah parolee data set contains 37 independent variables. This implies that there are more than 137 billion possible model specifications, or exactly 2^{37} possible models. In order to give a sense of the magnitude of this number in terms of computation time, consider that there are approximately 31.5 billion seconds in 1,000 years. Thus, it is impossible for any individual to compute and visually inspect 137 billion OLS linear regressions in a lifetime. Stepwise methods for finding the best model, whether bottom-up or top-down, are extremely risky, if not altogether unreliable. At any step in the process, the elimination of a single variable may dramatically change the t statistics and magnitudes for the remaining coefficients, thereby creating a false impression of which variables belong to the best model. A coefficient may have a low t statistic for only 1% of the entire set of models, but if the stepwise procedure begins inside that 1% the variable will likely be discarded from all future consideration. BMA effectively addresses this problem by exhaustively searching for the best model specifications and using those models to assign probabilities to the independent variables indicating their likelihood of being members of the best model.

As its name suggests, BMA is a technique that examines a large set of the most-likely models and estimates averaged coefficients based on the frequency with which the variables are selected along with the magnitudes of the coefficients in each case. Model combining and model averaging were studied as early as the 1960s, but it was not until 1978 that Edward Leamer fully elaborated the ideas that would eventually develop into Bayesian Model Averaging (Hoeting, Madigan, Raftery, & Volinsky, 1999). BMA became practical in the mid-1990s only after several theoretical and computational advances. Raftery, Madigan, and Hoeting (1997) and Hoeting et al. (1999) provide theoretical discussions of BMA and the following theoretical overview is based on these two sources.

The fundamental concept in BMA is the calculation of probability distributions for each of the coefficients based the probability distributions for the coefficients in each of the selected models and the probabilities associated with each model's being the best model. Let $\mathcal{M}$ represent the set of k possible models: $\mathcal{M} = \{M_1, \ldots, M_k\}$. If Δ is some quantity of interest, such as a coefficient for a linear regression, the posterior distribution of Δ given the data, D, is defined as

$$\Pr(\Delta|D) = \sum_{k=1}^{K} \Pr(\Delta|D, M_k)\Pr(M_k|D). \tag{5.1}$$

The posterior distribution in Equation (5.1) represents a weighted average of the probability distributions for Δ for each model k, where the weight is the probability that model k is the true model. It is assumed that the true model is encompassed by the k models so that $\sum_k \Pr(M_k|D) = 1$. The posterior model probability for model k is defined as

$$\Pr(M_k|D) = \frac{\Pr(D|M_k)\Pr(M_k)}{\sum_{l=1}^{K} \Pr(D|M_l)\Pr(M_l)}. \tag{5.2}$$

Equation (5.2) is an application of Bayes's Theorem to the issue of model uncertainty. The posterior probability for model k will be proportional to the prior distribution for model k, $\Pr(M_k)$, times the integrated likelihood function for model k, which is given in Equation (5.3).

$$\Pr(D|M_k) = \int \Pr(D|\theta_k, M_k)\Pr(\theta_k|M_k)d\theta_k \tag{5.3}$$

For large data sets with many independent variables, a completely exhaustive calculation of Equation (5.1) for all variables is not practical. Instead, a smaller subset of models having the highest posterior model probabilities is selected. There are three methods generally used to determine the subset of the best models: Occam's window, the leaps and bounds algorithm, and Markov chain Monte Carlo model composition. The BMA models for the Utah parolee data were fitted using the statistical software R. Model selection was performed using the leaps and bounds algorithm, so only this method is described.

The leaps and bounds algorithm was developed by Furnival and Wilson (1974) specifically for addressing the model selection problem when given large data sets. The goal is to identify the best independent variables in terms of their predictive accuracy on the basis of having the smallest residual sum of squares (RSS). The method relies upon the fact that if A is a model where the variables are a subset of the variables in model B, then $\text{RSS}(B) \leq \text{RSS}(A)$. An appropriately chosen branching method can be used to exhaust all of the possible models given an initial set of variables. An initial lower bound can be found by randomly selecting a single variable and calculating its RSS. A first wave of subsets (i.e., model specifications) is selected and the RSS is calculated for each subset. For those subsets with an RSS greater than the initial lower bound, they can be ignored because none of the variables in that subset can have a lower RSS than that for the initial variable. For those subsets with an RSS lower than the initial bound, the subset is branched to form additional subsets and the RSS is calculated for each new subset. Once a variable is found with a lower RSS than the initial lower bound, this lower RSS becomes the new lower bound. This process greatly reduces the time needed to find the best model specifications by skipping over those sets of model specifications that cannot possibly have variables with the lowest RSS.

In addition to the problem of finding the subset of the best models, another difficulty in applying BMA is integrating the likelihood function in (5.3). This can be approximated using the Bayesian Information Criterion (BIC). For linear regressions, the BIC function is defined as

$$\mathrm{BIC}_k = n \log(1 - R_k^2) + p_k \log n. \tag{5.4}$$

In equation (5.4), R_k^2 is the R-squared for the kth model, p_k is the number of regressors in the kth model, and n is the number of observations in the data set. The higher the R_k^2 score, the lower is the BIC score. Naturally, those models with the lowest BIC scores are considered the best models. A low BIC score, or a high R^2 score, means that the model fits the data relatively well, which will produce a relatively high $\Pr(D|M_k)$ in Equation (5.3). This, in turn, implies that the posterior model probability, $\Pr(M_k|D)$ in Equation (5.2), will likewise be relatively high. Thus, there will be a natural correspondence between how well a particular model predicts the data and its posterior model probability of being the best model.

5.3.2 Results of the analysis

BMA was applied to the Utah parolee data set to predict returns to prison using 37 independent variables. The averaging process was applied to an underlying linear probability model structure. Although prior probabilities can be assigned to each of the variables, no prior probabilities were assumed. Thus, the data were left to "speak for themselves." A few variables were removed before running the analysis. The hours worked per week variable was removed because of high collinearity with the employment variable ($r = .9$). The race variable "White" was removed to avoid perfect collinearity, so all race coefficients can be interpreted as relative to Whites. Likewise, the crime type "other" was excluded to avoid perfect collinearity, so all crime type coefficients can be interpreted as relative to other crimes. The results of the BMA analysis are presented in Table 5.2.

In Table 5.2, "$p \neq 0$" denotes the probability that the coefficient is not zero, "EV" represents the expected value of the coefficient, and "SD" is the standard deviation for the coefficient. The leaps and bounds algorithm selected the 74 best models, five of which are produced in the table. The five models in the table are sorted from smallest to highest BIC score, which corresponds to the models being "most likely correct" to "least likely correct." This correspondence is also reflected in the posterior model probabilities given in the last line in Table 5.2 for the five best models.

Turning to the results for each variable, the discussion separates the variables into the categories of economic and noneconomic variables. Beginning with the economic variables, the single-most important variable for predicting recidivism is employment. With a posterior probability of 100, the employment variable certainly belongs to the best model of recidivism. The absolute value of the coefficient is the second largest among all coefficient, implying it has a large effect on recidivism. If a parolee is employed, it reduces the probability of returning to prison by 20 percentage points relative to an unemployed parolee. The second most important economic variable is restitution. With a posterior probability of 61.4%, this variable is very likely a member of the best model. If a parolee pays restitution, the

Table 5.2. Bayesian Model Averaging Coefficient Estimation

74 models were selected
Best 5 models (cumulative posterior probability = 0.1907):

	$p \neq 0$	EV	SD	M_1	M_2	M_3	M_4	M_5
intercept	100.0	.643	.134	.662	.709	.684	.725	.678
age	79.0	−.005	.003	−.007	−.007	−.007	−.007	−.007
male	.0	.000	.000	.	.	.	.	.
priorinc	92.8	.045	.020	.053	.049	.049	.048	.052
hispanic	.0	.000	.000	.	.	.	.	.
pacisland	.0	.000	.000	.	.	.	.	.
natamer	.0	.000	.000	.	.	.	.	.
aframer	56.5	.129	.132	.229	.225	.230	.	.
asian	57.2	−.287	.289	−.518	−.488	.	−.501	−.532
drug	.0	.000	.000	.	.	.	.	.
driving	.0	.000	.000	.	.	.	.	.
murder	.0	.000	.000	.	.	.	.	.
person	8.5	−.011	.040	.	.	.	.	.
property	33.0	.043	.067	.	.129	.133	.134	.
sexcrime	3.6	−.006	.031	.	.	.	.	.
weapons	.0	.000	.000	.	.	.	.	.
highsch	2.5	−.003	.018	.	.	.	.	.
ged	.0	.000	.000	.	.	.	.	.
college	.0	.000	.000	.	.	.	.	.
vocation	.0	.000	.000	.	.	.	.	.
pednone	.0	.000	.000	.	.	.	.	.
pedged	14.0	−.016	.043	.	.	.	.	.
pedhs	.0	.000	.000	.	.	.	.	.
pedpost	.0	.000	.000	.	.	.	.	.
employed	100.0	−.200	.050	−.200	−.193	−.191	−.201	−.208
wage	.0	.000	.000	.	.	.	.	.
healthben	.0	.000	.000	.	.	.	.	.
managejob	.0	.000	.000	.	.	.	.	.
buildjob	.0	.000	.000	.	.	.	.	.
salesjob	.0	.000	.000	.	.	.	.	.
officejob	.0	.000	.000	.	.	.	.	.
constrjob	.0	.000	.000	.	.	.	.	.
installjob	.0	.000	.000	.	.	.	.	.
prodjob	.0	.000	.000	.	.	.	.	.
transjob	.0	.000	.000	.	.	.	.	.
restitution	61.4	.074	.068	.118	.	.	.	.121
childsup	38.2	.042	.061	.	.	.	.	.
owncar	27.4	−.030	.054	.	.	.	.	.
Number of Variables				6	6	5	5	5
R^2				.109	.108	.097	.097	.097
BIC				-20.948	-20.749	-20.625	-20.569	-20.535
Posterior probability				.043	.039	.037	.036	.035

probability of returning to prison increases 7.4 percentage points. The child support variable is the only other economic variable with strong results. The posterior probability of 38.2% indicates that, while not dominant, the child support variable has a significant role to play in predicting recidivism. The probability of returning to prison increases by 4.2 percentage points when a parolee must pay child support. The wage variable appears unimportant and this is because it is highly collinear with the employment variable ($r = .79$). In the presence of high collinearity, BMA typically selects one variable and ignores the other. Health benefits and occupational type do not appear to be important predictors of recidivism.

The educational variables do not appear to be strong predictors of recidivism. One possible explanation is that because offenders must report felonies on employment applications, the effects of education on employment status and the wage rate are suppressed. Of all of the educational variables, only the educational attainment variable for high school diploma and the prison education variable for GED preparation have some importance. In both cases, they slightly reduce the probability of returning to prison.

As for the noneconomic variables, the number of prior incarcerations is the strongest predictor with a posterior probability of 92.8%. For each prior incarceration, a parolee is 4.5 percentage points more likely to return to prison. With a posterior probability of 79%, age at the time of the survey is the next strongest predictor of recidivism. For each additional year in age, the probability of returning to prison drops by .5 percentage points. Two race variables were also important predictors. The Asian variable has a posterior probability of 57.2% and the African American variable has a posterior probability of 56.5%, which makes both of these variables important. The probability of an Asian parolee returning to prison is 28.7 percentage points lower than for a White parolee, while the probability of returning to prison for an African American parolee is 12.9 percentage points higher than a White parolee. The last variable of importance is the type of crime most recently committed before the survey was administered. The property crime variable has a 33% probability of being in the best model. If a parolee's most recent conviction was for a property crime, the probability of returning to prison is 4.3 percentage points higher in comparison to those in the "other crime" category.

The method by which the posterior probabilities for the coefficients are determined is graphically represented in Figure 5.2. All 74 of the best models are listed along the horizontal axis in order of "most likely correct" to "least likely correct." For each model, the variables included in that model are indicated by the black bars. The determination of the posterior probability for a coefficient will depend on the number of times the variable is used weighted by the model's ranking in which it is used.

The usefulness of Figure 5.2. lies in the fact that the values found under the "$p \neq 0$" column of Table 5.2 are presented in such a way that the relative importance of each variable can be easily visualized and understood. For instance, employment status is the most important predictor of recidivism and, with a posterior probability of 100, it is certainly a member of the best model. Correspondingly, a solid black bar traverses the entire range of the 74 best models indicating its inclusion in every model. The next two strongest predictors of recidivism are prior incarcerations and age, and the black bars representing their inclusion in the models reveal only a small number of gaps where they were not included. The restitution variable is the fourth strongest predictor and its importance can be seen by its inclusion in the majority of the models. Likewise, the importance of the child support variable and the Asian and African American race variables can be immediately perceived.

A few comments regarding the fit of the BMA model deserve mention. Using the data to evaluate the fit of the BMA model, the model correctly classifies 64.8% of parolees. The model predicts that 149 of the parolees will return to prison, whereas 221 actually returned.

5.3.3 A 10-fold cross-validation of three recidivism models

Of central importance to this dissertation is the determination of the relative importance of economic variables in predicting recidivism versus noneconomic variables. Information regarding noneconomic variables, such as age, race, and gender, is collected as a matter of policy by corrections departments and, as a result, is

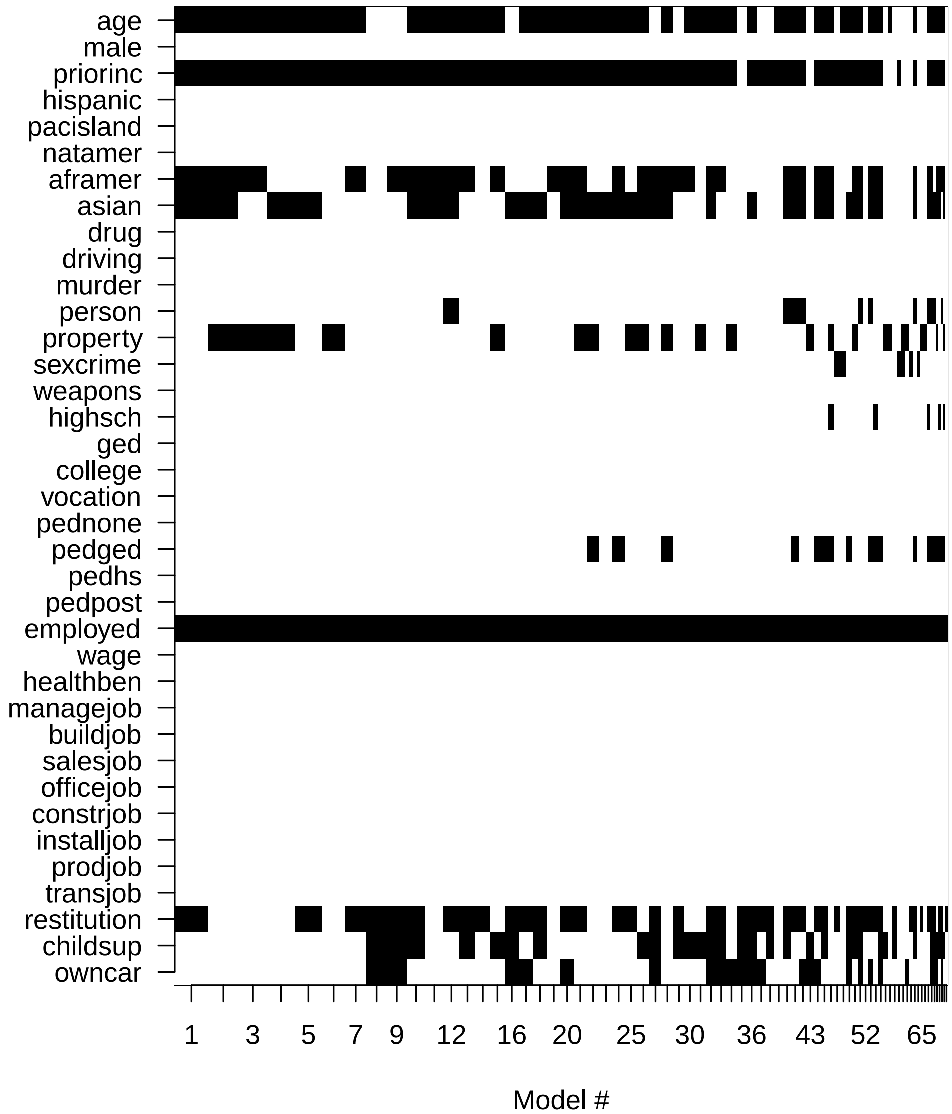

Figure 5.2. Models Selected by Bayesian Model Averaging

relatively inexpensive to obtain. Economic information, on the other hand, is more costly to acquire. Thus, the collection of economic data on parolees and probationers is only justified in so far as it leads to a better understanding of those factors that increase the risk of recidivism.

In order to determine the relative importance of economic variables in predicting recidivism, a 10-fold cross-validation of three models is performed. The three models used in this experiment are referred to as the Sociological Model, the Economic Model, and the Total Model. The Sociological Model includes only those variables that are sociological in nature and can be easily obtained from any corrections department. These variables include age, gender, race, prior incarcerations, and most recent crime committed. The Economic Model is based solely on the variables for employment, restitution, and child support. The Total Model uses all of the variables. The results for the three models are found in Table 5.3.

The 10-fold cross-validation was performed by randomly sorting the data and dividing it into 10 more-or-less equal subsets. Each subset contained 51 observations and two subsets had four observations in common. For the three models, 90% of the data was used as training data for estimating the coefficients, then the model was used to predict the remaining 10%. This was performed systematically so that predictions were made for each of the 10 subsets based on the other 90% of the data. Thus, every observation served as training data and testing data for all three models. For each model, BMA was used to estimate the coefficients for a linear probability model using only those variables included in the model and excluding all other variables. Naturally, the Total Model uses all of the variables. Two measures were used to evaluate predictive accuracy. The first measure is the number of correct classifications over the total number of 51 predictions, which essentially produces a ratio of successes. The second measure is the mean square error, which measures how close the predicted value is to the actual value.

Looking across the folds, each model performs relatively better than some other model for a given fold and relatively worse in another fold. The last column gives the overall results as averages for the 10 folds for each model. The Economic Model, which used only the three variables of employment, restitution, and child support, correctly classified parolees slightly better than the Sociological Model. When all of the information was used, the Total Model produced a correct classification rate approximately 2.5 percentage points higher than the other two models. The Economic Model had the lowest mean square error, which implies that the predicted and actual values were closer on average than for the other two models. The results

Table 5.3. Results of Three 10-Fold Cross-Validations

Correct Classifications											
Fold	1	2	3	4	5	6	7	8	9	10	Overall
Sociological Model	.5098	.5686	.6078	.5882	.6471	.6078	.6275	.5098	.5882	.5686	.5824
Economic Model	.5490	.5686	.5490	.6471	.6078	.6275	.6078	.5098	.6275	.5686	.5863
Total Model	.5490	.6078	.6078	.7451	.6667	.5882	.6275	.5294	.5882	.5882	.6098
Mean Square Error											
Fold	1	2	3	4	5	6	7	8	9	10	Overall
Sociological Model	.2396	.2377	.2474	.2212	.2324	.2277	.2427	.2619	.2380	.2420	.2391
Economic Model	.2430	.2435	.2404	.2204	.2237	.2484	.2242	.2588	.2175	.2334	.2353
Total Model	.2513	.2418	.2574	.2171	.2232	.2350	.2259	.2621	.2339	.2537	.2402

suggest that economic variables are just slightly better predictors of recidivism than sociological variables. Furthermore, when both economic and sociological variables are used together, the correct classification rate can be improved significantly.

5.4 Classification and Regression Trees

While the use of trees for classification and regression dates back to the 1960s, the first complete, formal treatment of tree methodology was produced by Breiman, Friedman, Olshen, and Stone (1984). The CART method encompasses two types of models. A classification tree is used to model an equation where the response variable is a categorical variable and a regression tree is used for a numerical response variable. As the interest here lies with predicting a binary categorical variable (i.e., return to prison), the models developed here are classification trees.

Tree methodology has several advantages over other methods of statistical classification. Because tree models are easy to interpret, they can be implemented within practical settings by individuals with essentially no statistical background. Furthermore, the informational and computational requirements for classifying observations through the use of a tree are minimal. A regression equation requires knowledge of all values of the variables and often involves considerable computation in order to determine the class membership of an observation. A tree, on the other hand, can typically classify a large number of observations using only a few variables and, in some cases, the necessary computations are reduced to determining the values of binary variables. More importantly, the tree methodology offers an effective method for treating nonhomogeneous relationships. A classification tree is essentially a series of conditional statements and these conditional statements divide observations into subsets corresponding to the different qualities of the nonhomogeneous relationship (e.g., the observations determining a quadratic relationship can be divided into those exhibiting a positive relationship and those exhibiting a negative relationship). Not only can a classification tree reveal information about nonhomogeneous relationships within the data simply through its structure, but it will tend to better fit the data and produce more accurate predictions when taking account of nonhomogeneous relationships as compared to linear classification methods.

The tree models found in this dissertation were produced using the statistical software R. The characterization of the tree construction process follows Breiman et al. (1984). Forming a classification tree involves three aspects: determining the splitting rules, deciding when to stop splitting nodes, and assigning terminal nodes to classes. When selecting a splitting rule for a node, the goal is to choose an independent variable and a split point such that the observations are separated into two subsets, each of which is *purer* than the predecessor set. The notion of *purity* is defined formally in terms of an impurity function $i(t)$, where $i(\cdot)$ is a nonnegative function measuring the impurity at node t. The function $i(t)$ will be at its largest value when all of the class types are mixed together and the function will be zero for some node when its only members are of one class. In terms of the prisoner recidivism problem, the data set of 506 parolees contains 43.7% who returned to prison and 56.3% who did not. For this initial node, the impurity function $i(t)$ would be at its maximum. If it were possible to completely separate those who returned from those who did not, the function $i(t)$ would be zero at each of the two terminal nodes containing the 221 who returned and the 285 who did not.

The impurity function is ultimately used to determine the best independent variable and the best split point along the range of that variable for a splitting rule at a particular node. Large data sets can produce millions of possible splitting rule and some criterion must be used to find the best rule. Numerical variables can theoretically produce an infinite number of splitting rules, but for practical purposes only several thousand split points might be considered. Binary categorical variables will only have one possible split point. Let $\mathcal{S}$ denote the set of all possible splits corresponding to all of the independent variables and all of their split points, where s denotes an individual split such that $s \in \mathcal{S}$. Breiman et al. (1984) define the goodness of the split s as the decrease in impurity measured by the function

$$\Delta i(s,t) = i(t) - p_L \cdot i(t_L) - p_R \cdot i(t_R), \tag{5.5}$$

where t is the initial node, t_L and t_R are the left and right successor nodes, p_L is the proportion of observations in node t_L taken from t, and p_R is the proportion of observations in node t_R taken from t. Each split s will have a corresponding measure of goodness. Finally, the best splitting rule s^* for some node t is determined as the one that produces the greatest reduction in impurity and is defined as

$$\Delta i(s^*, t) = \max_{s \in \mathcal{S}} \Delta i(s, t). \tag{5.6}$$

Given the nature of the splitting method, those splits that occur earlier in the tree construction process usually indicate variables that are stronger predictors of the dependent variable. Once the best split s^* has been determined for the initial node, the splitting procedure is performed anew on each of the successor nodes, and so on. The way in which the tree methodology treats nonhomogeneous relationships can be understood in relationship to the splitting method described above. Assuming two independent variables x_1 and x_2, it might be that no clear relationships between either x_1 or x_2 and the dependent variable y can be perceived when examining linear regression output or even scatterplots of the data. When constructing a tree, if a split occurs using the variable x_1, the resulting subsets in the successor nodes may exhibit very strong relationships between x_2 and y. This conditional approach to modeling the data essentially reveals relationships that may exist only within subsets and not at the level of the entire data set. Because linear regression models only operate at the level of the entire data set, they cannot detect relationships that occur only within subsets of the data (i.e., nonhomogeneous relationships).

Turning to the Utah parolee data, two classification trees were modeled on the data. Both are interpreted as follows. Observations that satisfy the splitting rule at a given node are classified to the left branch and those that do not satisfy the rule are classified to the right. The vertical length of a branch connecting two nodes indicates the importance of the variable at the predecessor node. The longer the branch, the greater is the decrease in impurity. The probability for each terminal node is calculated as the number of parolees that returned to prison divided by the total number of parolees in the terminal node. If the probability at the terminal node is greater than or equal to .5, then all parolees classified to that terminal node are predicted to return to prison within three years. Nodes with probabilities less than .5 determine sets of parolees that are predicted not to return to prison.

The best tree model, CART Model 1, is presented in Figure 5.3. The tree was constructed using the same data set as was used for the BMA model with the exception that the wage variable was removed. The data set contains 506 observations and 36 independent variables. The model predicts that a total of 194 of the 506 parolees from the data set will return to prison within three years producing a recidivism rate of 38.34%. Regarding the model fit, the tree correctly classifies 72.53% of the observations. This is the best correct classification rate for any of the models that were estimated using the Utah parolee data in this dissertation.

The selection of variables used for splitting rules in CART Model 1 indicates the importance of these variables to predicting recidivism and the results are generally consistent with the BMA results. The BMA model indicated that employment, restitution, age at release, and prior incarcerations were the most important predictors of recidivism and all of these variables are present throughout the tree. The lengths of the branches emanating from the employment and restitution nodes indicate that these two economic variables are among the strongest predictors of recidivism. Employment status is the splitting rule for the initial node, which shows its importance in separating those predicted to return to prison from those predicted not to return. With 383 employed parolees classified to the restitution node, the restitution splitting rule acts upon the largest subset of the data. Of the 383 employed parolees classified to the restitution node, 201 were making restitution payments, which represents 79.4% of all parolees making restitution payments. It is interesting to note that while the educational attainment variables GED and high school diploma and the sex offense crime variable were not considered important in the BMA model, they play relatively important roles in the tree model. This indicates that these variables exhibit nonhomogeneous relationships throughout the

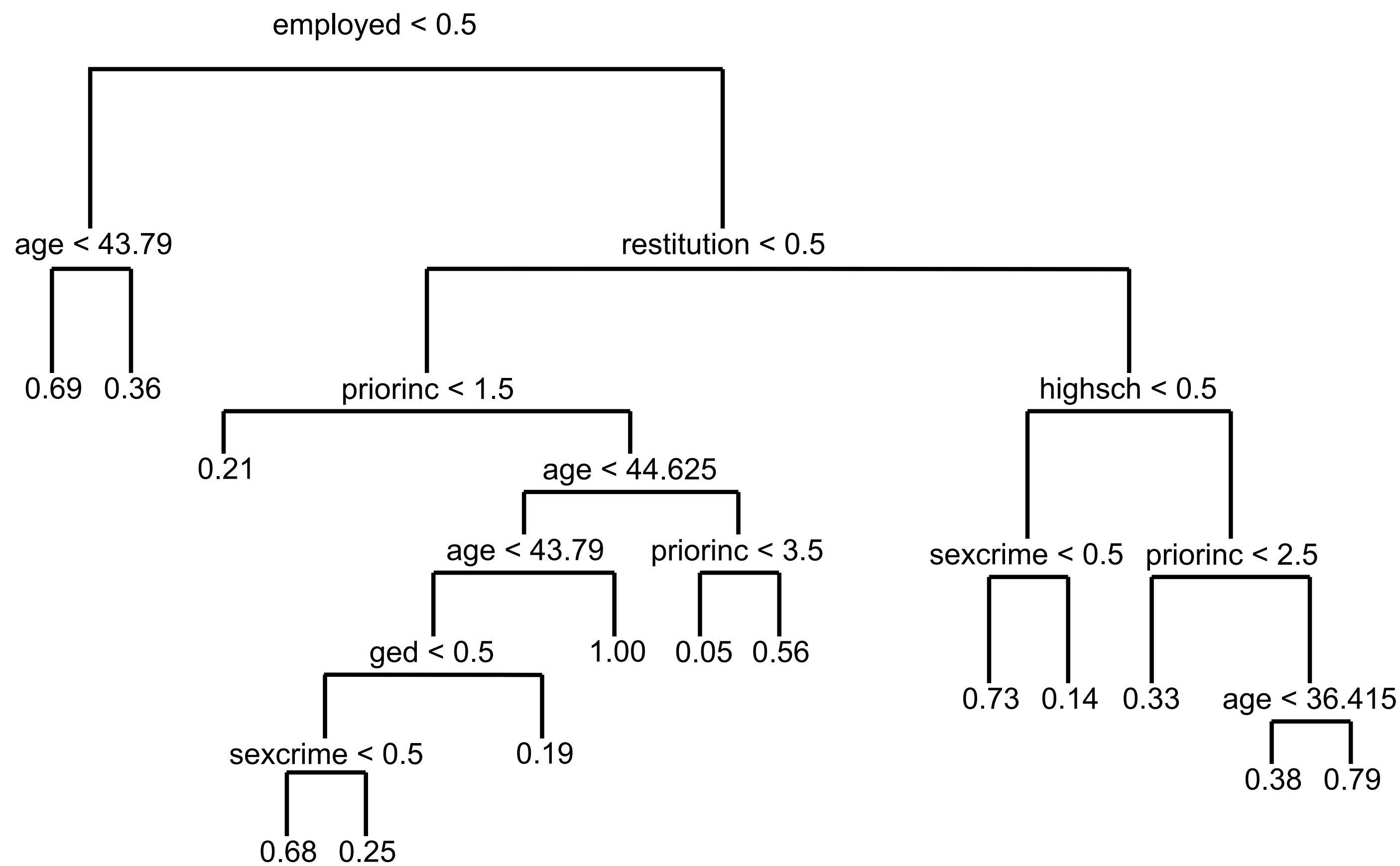

Figure 5.3. CART Model 1

range of the entire data set and linear classification methods that operate at the level of the entire data set cannot detect their relationship to recidivism. When conditionally separated into subsets, all of these variables possess relatively strong relationships to recidivism.

A second tree model, CART Model 2, is presented in Figure 5.4. The high collinearity between the employment and wage variables ($r = .79$) led the tree modeling process to choose only one of these variables and ignore the other. CART Model 2 was fitted using the same data set of 506 observations and a total of 36 independent variables, but the wage variable was included and the employment variable was excluded. As expected, the high collinearity between the wage and employment variables led to the substitution of the wage variable for the employment variable at exactly the same location in both tree models.

The two tree models share many similarities. Consistent with the BMA model and CART Model 1, the variables wage/employment, restitution, age at release, and prior incarcerations all play important roles in CART Model 2. The restitution variable is still of considerable importance in the second tree model as 257 total observations are classified to the restitution node for further branching. However, beyond the initial node, there are several structural differences between the two tree models. The GED, high school diploma, and sex offense variables of CART Model 1 are not present in CART Model 2. Instead, three new variables are present. CART Model 2 includes the variables for postsecondary prison education, sales as an occupational type, and drug offense as the most recent crime committed by the parolee. In the BMA model, none of these variables were important. The presence of these variables in the tree model implies that these variables exhibit nonhomogeneous relationships and are important predictors of recidivism only within subsets of the entire data set.

The model fit for the second tree model is nearly as good as for the first tree model. CART Model 2 correctly classified 72.13% of the observations. The model predicts that 230 of the 506 parolees will return to prison within three years, which is a recidivism rate of 45.5%. Even though the second tree model predicts a recidivism

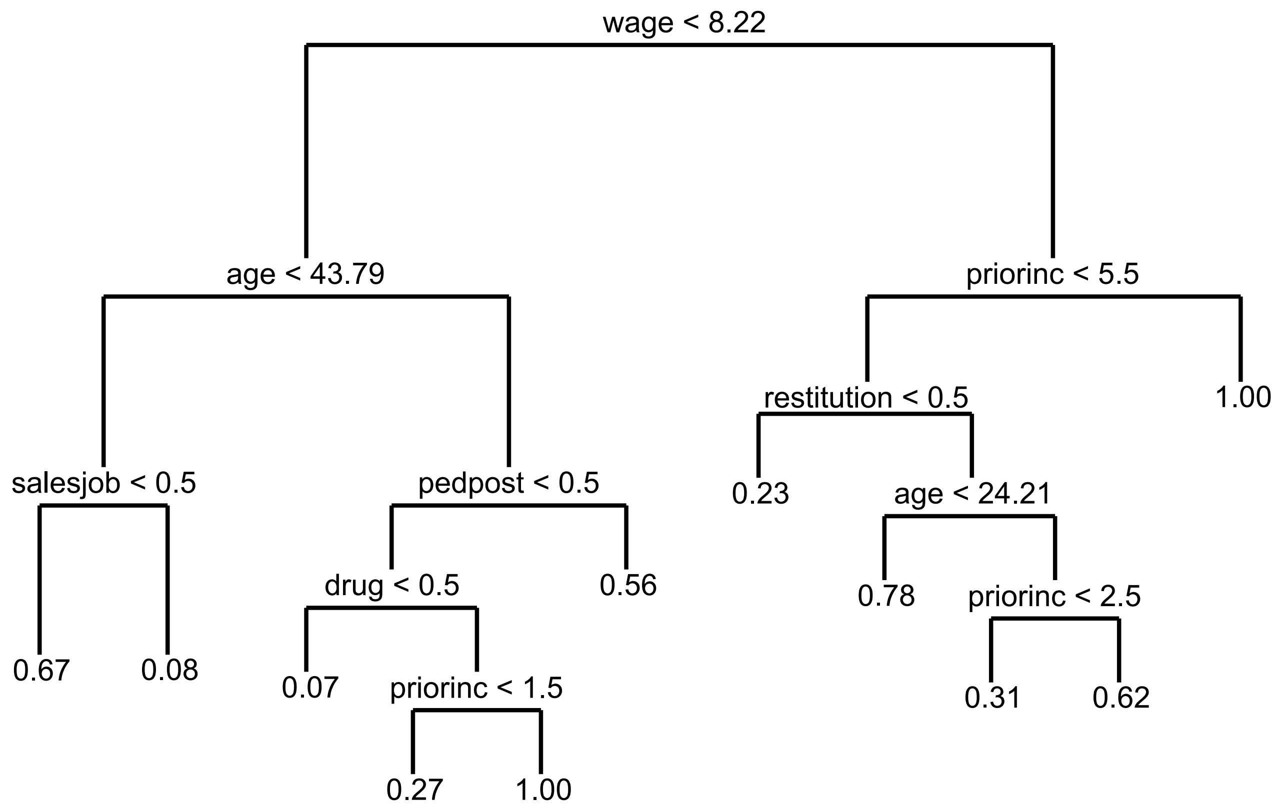

Figure 5.4. CART Model 2

rate that is closer to the actual recidivism rate than the first model, it overestimates the recidivism rate. On this basis, CART Model 1 might be viewed as preferable for the purpose of risk assessment because arguably a greater injustice occurs from incorrectly classifying nonrecidivists as recidivists than from incorrectly classifying recidivists as nonrecidivists.

5.5 A linear probability model

The last model presented is an OLS linear probability model estimated using the Utah parolee data. The inclusion of this model is primarily for comparative purposes. As mentioned previously, over 137 billion different model specifications can be produced from the set of 37 independent variables used to analyze the Utah parolee data. With such a large number of possible choices, the question arises as to how the particular model specification chosen here was settled upon as being the best model. The answer, in this case, is simple: The variable selection was determined by the BMA analysis. All of the variables included in the regression correspond to the variables that had positive posterior probabilities in the BMA output in Table 5.2. Besides providing a more familiar set of statistical results for the parolee data, the OLS regression is useful for demonstrating the differences between this classical technique and the two Bayesian techniques.

The OLS regression results for the linear probability model are reported in Table 5.4. The model correctly classified 66.8% of the data, which is 2 percentage points greater than the BMA model and 3.7 percentage points less than CART Model 1. Using the parolee data, it predicts 183 will return to prison, which implies a predicted recidivism rate of approximately 36%.

A comparison of the OLS regression results with the results from the BMA and CART analyses reveals several advantages of the Bayesian methods over the classical method. First and foremost, both the BMA and CART methods provide solutions to the model specification problem, whereas the OLS linear regression method, in and of itself, is silent on this issue. Moreover, using the results of an OLS linear regression to determine model specification may lead to the selection of the

Table 5.4. Linear Probability Model

	Estimate	Std. Error	t value
Intercept	.760225	.107300	7.085
age	−.005612	.002222	−2.526
priorinc	.046091	.014975	3.078
aframer	.216870	.088010	2.464
asian	−.442395	.192963	−2.293
person	−.128181	.062540	−2.050
property	.034771	.054745	0.635
sexcrime	−.067588	.060546	−1.116
highsch	−.080350	.047700	−1.684
pedged	−.133274	.051165	−2.605
employed	−.190627	.049320	−3.865
restitution	.094169	.044958	2.095
childsup	.092202	.044803	2.058
owncar	−.088958	.043912	−2.026

wrong variables. Referring to Table 5.4, a researcher using a stepwise elimination approach may decide to reject the property crime, sex crime, and possibly the high school diploma variables due to their relatively low t statistics. However, both the BMA and CART methods indicate that these variable have a role to play in predicting recidivism. In the BMA results in Table 5.2, the posterior probabilities associated with the property crime, sex crime, and high school diploma variables were 33%, 3.6%, and 3.5%, respectively. While these low probabilities indicate that these variables are somewhat weak predictors, it seems better to include them rather than omit them completely. CART Model 1 adds further evidence that at least two of these variables play a significant role in prediction. The high school diploma variable is used to separate out 201 observations and the two occurrences of the sex crime variable together separate out 93 observations, which indicates that these variables are relevant for predicting recidivism. In their different ways, the CART and BMA approaches find these variables important while the OLS linear regression model does not.

Another issue than can be illustrated by way of these comparisons concerns the estimated magnitude of the coefficients. The OLS estimates in Table 5.4 for age, prior incarcerations, and employment status are very similar to the BMA estimates of -.005, .045, and -.2, respectively. This results from the facts that the BMA posterior probabilities are very high for these variables and that the OLS estimation procedure effectively assumes these variables are certainly members of the best model. As for the other variables, the magnitudes of the OLS coefficients are quite different as compared to the BMA expected values. BMA adjusts the size of the coefficients to reflect the uncertainty expressed in the posterior probabilities for each variable. Consequently, OLS produces coefficients that are too large because it assumes the variables all belong to the best model. In a sense, OLS estimation leads to a different type of over-fitting problem, one stemming from its inability to incorporate uncertainty about the variables into the estimation procedure. Similar to the problem of over-fitting as the term is used in its ordinary sense, the OLS model can be expected to perform worse on out-of-sample predictions in comparison to the BMA model.

5.6 A summary of the results

The analyses reveal that there are seven variables that are of particular importance to predicting recidivism. Among the economic variables, employment status, restitution payments, and child support payments are the most important. As for the sociological variables, age, prior incarcerations, and the Asian and African American race variables were the strongest predictors. Together, the BMA and CART models clearly show that the four most powerful predictors are employment, age, prior incarcerations, and restitution.

In comparison with past recidivism studies, the importance of the variables age, race, prior incarcerations, and crime type as found in the Utah parolee data set is consistent with past results. The gender, education, and Hispanic race variables were found to be unimportant in the Utah parolee data set, which generally conforms to the results from past studies. Even though wages were removed from the BMA model due to high collinearity with employment, wages were found to be significantly related to recidivism in the Utah parolee data set, which is consistent with the past studies. However, the relationship between employment status and recidivism was found to be of uncertain significance in previous studies and it is the single-most important variable in the parolee data set. Restitution payments, which exhibited a somewhat weak negative relationship with recidivism in earlier research, showed a strong positive relationship. Finally, job type was found to be related to recidivism in past studies, but no significant relationship was found in the Utah data set.

The 10-fold cross-validation showed that economic variables are just slightly better predictors of recidivism than sociological variables. The Total Model, composed of both economic and sociological variables, performed even better, indicating that the best model would incorporate a mix of these variables.

Regarding model fit, the BMA model correctly classified 64.8% of Utah parolees, while the OLS linear probability model correctly classified 66.8%. The tree methodology developed models that fit the data much better than either the BMA or OLS models. CART Model 1 correctly classified 72.5% of the parolees and CART Model 2 correctly classified 72.1%.

CHAPTER 6

CONCLUSIONS

This dissertation began by examining the problems of the ever-increasing incarceration rate in the United States and the concomitant rise in corrections spending. The growing share of corrections spending relative to direct general expenditure is placing a heavy burden on state budgets throughout the country. The solution to this problem considered in this study is to reduce recidivism. Specifically, the focus has been to determine the importance of economic factors in predicting recidivism. The analysis in Chapter 5 demonstrated that economic factors have a strong impact on recidivism. Attention now turns to an examination of the policy implications resulting from the previous analysis. The first section of this chapter examines policies concerning employment, restitution payments, and child support payments and includes policy recommendations for these economic variables.

Even when a particular policy decision appears clearly beneficial to society, such a policy may never be implemented due to the nature of the political system. The second section of this chapter contains a brief discussion of the political difficulties involved with changing criminal justice policy, even in the presence of compensations schemes that could theoretically benefit all parties.

The final section serves as a proposal for future recidivism research on several topics. Some topics involve exploring other aspects of the Utah parolee data set that were beyond the scope of this dissertation. Other topics concern variables other than those for which information was collected from the Utah parolees. Research into these variables is of considerable importance because criminal justice policy needs to be based on an understanding of how various sanctions influence the likelihood that released prisoners will engage in future criminal activity.

6.1 The policy implications

The criminal justice policy implications of the preceding statistical analyses are considered in this section. The focus centers upon policies involving the economic variables of employment, restitution, and child support. The Utah parolee data set and the models developed in the previous chapter are used to examine changes in the predicted level of recidivism resulting from changes in these economic variables. Policy recommendations are then formed by considering the changes in corrections costs implied by the changes in predicted recidivism associated with each of the possible policy choices.

In practice, calculating all of the explicit and implicit costs and benefits arising from the change in one variable can be extremely difficult. To calculate the cost of a single instance of recidivism based on a new crime requires knowing the cost to victims, law enforcement costs, adjudication costs, and corrections costs. These are only the explicit costs. A complete cost assessment would include the quantities of goods and services the offender could have produced if not incarcerated, the reallocation of resources to the production of crime prevention goods resulting from the increased probability of crime, the increases in insurance costs associated with crime-related property damages, and so forth. While the complexity of a complete cost-benefit analysis associated with the commission of a crime is acknowledged, the approach taken here is, by necessity, greatly simplified. The cost of recidivism will be considered solely in terms of the incarceration cost. This will produce a cost that is a significant underestimate of the true cost of recidivism because law enforcement and adjudication costs are ignored. If policy recommendations can be formed when using an underestimated cost figure, this underestimated cost will be sufficient for developing criminal justice policy recommendations. However, there is one case where the underestimated cost is not sufficient for making a strong policy recommendation with respect to one of the economic variables.

The policy recommendations developed here are largely intended for the consideration of criminal justice policy in Utah. The parolee data used to construct the models and the corrections cost figures are based on information specific to Utah.

Due to the unavailability of some types of information, estimates based on data from other state-specific or nationwide studies were used. Even though the policy discussion is more directly focused upon Utah, it could be applied to any other state or the nation as a whole by adjusting the cost information accordingly.

In order to derive any policy recommendations, some assumptions must be made. The assumption most liable to critical scrutiny is that the cost of incarcerating an individual is a socially less-desirable use of tax dollars than, say, the cost of education or the cost of building a public hospital. In response to the view that incarceration may be desirable if it improves public safety, the argument does not properly apply to the recommendations made here. The recommendations are made on the basis of modifying the incentives faced by released prisoners with the intention of shaping their choice with respect to engaging in future criminal activity. If no crime occurs because a released prisoner has been able to successfully transition to the role of a law-abiding citizen, there is obviously no objectionable act requiring punishment. Therefore, an instance of recidivism that could have been prevented certainly appear to be a net loss to society because public safety is better served by eliminating the criminal activity altogether through modifying the behavior of released prisoners. When considering how recidivism can be reduced, the only variables that will be changed are the economic variables of employment, restitution, and child support. Otherwise, it is assumed that the level of law enforcement, sentencing practices, corrections programs, and essentially all other variables are held constant. These assumptions guarantee that the detection and punishment of crimes remain virtually the same, with the only exceptions being the three policy variables under consideration. Hence, no assumption is made for a decrease in the level of public safety with respect to criminal activity.

The incarceration cost of recidivism is estimated first. To determine the cost for a single instance of recidivism, the expected length of a return to prison needs to be calculated. Even though only 506 observations were used from the Utah parolee data set for the statistical analysis due to missing values, there were 648 observations for which complete incarceration information was available. Using the set of 648 parolees, there were a total of 409 returns to prison for 284 distinct individuals during the 3-year observation period. Of these 409 total returns, 316 were completed during the 3-year period, while 93 were still incarcerated at the end of the period. The 93 offenders who were still incarcerated at the end of the period represent censored data and these observations were excluded. For the 316 completed returns to prison, the average length of stay in prison was 241 days. The data set of 506 observations used for the statistical analyses produced a very similar expected length of stay. From the set of 506 observations, there were 187 completed returns to prison for an average length of 238 days. The expected length of 241 days will be used for the cost calculations because it comes from a larger sample and should be more accurate. In Utah, the incarceration cost for one offender per day is $79.63. Using the estimated length of stay in prison of 241 days, a single return to prison for either a new crime or a technical violation has an expected cost of approximately $19,190.83.

Employment, the single-most important predictor of recidivism as based on the previous analyses, is considered first. In a national survey based on a sample of 7,000 offenders in jails, 29% were unemployed in the 6 months prior to their incarceration (James, 2004). In the Utah parolee data set, 24% of parolees were unemployed at the time of the survey. Given that the unemployment rate in Utah has been historically lower than the unemployment rate for the United States, the rate derived from the Utah parolee data set appears consistent with the national estimate. Referring again to the national survey of jail inmates, it was also found that 60% had annual incomes of less than $12,000 in the year prior to incarceration (James, 2004). Utah parolees appear to have had higher average annual wages than the jail inmates. Only 32% of the Utah parolees had annual incomes less than $12,000 and half of them had annual incomes of $17,100 or less.

From the policy perspective, the interest lies in determining the value of a policy that can increase employment among parolees after their release from prison. The perspective taken here will be to consider a policy that guarantees full employment for all released prisoners. There are two components to the valuation of this employment policy: the value of employment itself and the value of reduced costs associated with reduced criminality. The value of employment should be included in the calculation because any unused resource represents an economic inefficiency. The value of employment will be based on the most conservative estimate possible. Assume that released prisoners are paid the federal minimum wage of $7.25 and work 2,080 hours per year for a total gross annual income of $15,080. This represents pure value creation from the economic perspective because a previously unused resource is now earning its value in the market. During a 3-year follow-up period, this represents $45,240 per released offender. Taking the 123 unemployed parolees from the data set of 506 observation and multiplying this by the 3-year amount produces a total of $5,564,520. This is the direct benefit from having these unemployed resources utilized.

The second component in the valuation of a full employment policy is the value associated with the reduction in predicted recidivism. The Utah parolee data was used along with the BMA model to predict how recidivism would change given that all parolees were employed. After changing the employment status variable so that everyone in the Utah parolee data set was employed and running it through the BMA model, predicted recidivism decreased by approximately 52%. Based on the 221 parolees that actually returned, this implies that 115 would not return to prison. Multiplying this figure by the expected corrections cost per instance of recidivism produces a corrections cost reduction of $2,206,945.45. As a final adjustment, the 3-year total income figures are reduced for the 106 parolees that return to prison on the assumption that all returns to prison are from this group who could not find employment without some form of assistance. Note that this is an extreme assumption that is not likely to occur, but it is made to create the most conservative estimate possible. The expected length of a return to prison is 241, which is approximately eight months. Thus, $10,053.33 in lost earnings due to incarceration must be subtracted for each return to prison. For the 106 parolees that returned, this amounts to $1,065,653.33. Adding the 3-year income figure to the reduction in corrections cost, and then subtracting the lost earnings due to 106 returns to prison produces a total of $6,705,812.12.

The amount of $6,705,812.12 represents the potential value to society in terms of employing unused resources and eliminating corrections costs from reduced criminality based on the assumption that a policy could guarantee full employment. However, no policy is costless. The purpose of developing this estimate is to create an upper bound for the cost of a full employment policy such that any policy with a total cost less than the upper bound that successfully creates full employment will produce a net benefit for society. If the total amount is divided by the 123 unemployed parolees in the sample, the value of employment amounts to $54,518.80 per parolee for the 3-year time frame. Any policy that has a 3-year cost less than this amount and could guarantee full employment would appear to be justified. Policies such at employer tax credits, job counseling services, or job training programs that totaled up to just under $18,172.93 annually per parolee would be beneficial for society, under the assumption that the parolee does, in fact, become employed. It should be remembered that the reduction in costs due to decreased recidivism is greatly underestimated because the figure ignores law enforcement and adjudication costs. Furthermore, it was assumed that the 106 parolees who were predicted to return even while being fully employed all came from the group that received some form of employment assistance under this policy. This is a worst-case scenario that would be unlikely to occur because a some portion of recidivism could be expected from those who found employment on their own. Therefore, the actual value to society of full employment for released prisoners is much higher than the estimate produced above.

Turning to restitution, the concern is to determine the corrections costs that arise from the imposition of restitution. Estimating the amount of restitution that parolees owe victims is problematic. No information was collected from the Utah parolees regarding the total amount owed, so this figure must be estimated by other means. McLean and Thompson (2007) found that the average amount of restitution owed by probationers in Arizona was approximately $3,500. The U.S. Department of Justice (1998) reports that from a survey of 32 counties the average restitution amount owed by probationers was \$3,368. From the Utah parolee data, those required to pay restitution paid an average of \$167.74 per month, which amounts to just over \$4,000 if payments are made for 2 years. The payment amount is typically determined so that the full amount of restitution is paid before the terms of parole or probation expire. It is unknown how long restitution payments are paid on average in Utah. The amount of \$3,500 will be taken as a likely approximation. It is estimated that only 54% of the full restitution amount is paid before the parole or probation period expires (U.S. Justice Department, 1998). Using this percentage, the amount of restitution that can be expected per parolee is \$1,890. Of the 506 observations in the Utah data set, 255 were required to pay restitution. Thus, the full amount of restitution expected from this sample of parolees is \$481,950.

The next step is to determine the incarceration cost associated with restitution payments. Once again, the Utah parolee data and the BMA model are used to predict the change in recidivism when restitution is eliminated. By removing restitution payments from the Utah parolee data set, the BMA model predicts that recidivism decreases by approximately 22%. Of the 221 individuals that actually returned to prison, 48 fewer are predicted to return when no restitution is imposed. This implies an incarceration cost reduction of roughly \$921,159.84.

Strictly speaking, restitution is a transfer payment that by itself does create net value. Restitution payments are imposed upon criminal offenders with the belief that it is just or right for the victim to be compensated for his or her loss due to the criminal behavior of the offender. While there may be good moral reasons for restitution payments, the policy recommendation here is based solely upon an examination of the costs. The expected reduction in incarceration costs of \$921,159.84 attributable to the elimination of restitution payments represents true value creation because crime would be reduced, public safety would remain constant, and corrections spending would fall. However, victims may not find the elimination of restitution payments palatable. Yet, in this case, the expected cost of incarceration far exceeds the expected amount of restitution, which implies that the government could theoretically pick up the tab for restitution on behalf of the released prisoners and still save $439,209.84 This figure is based only on the sample of 506 parolees. If the figure were estimated for the total parolee and probationer population of 16,000 individuals in Utah, the total cost savings would amount to approximately just over $13,888,000. Similar to the case of simply eliminating restitution payments altogether, a policy that requires the government to cover the costs of restitution would likely be perceived as unpalatable. Whatever the appearances may be, taxpayers could benefit, either in terms of tax reductions or having their tax dollars spent on education rather than prisons, from having government pay for restitution rather than released prisoners.

The last economic policy variable considered here is child support. As in the case of restitution, the estimation of the size of child support owed is problematic. Information on the total amount of child support owed in arrears was not collected from the Utah parolees, so it must be estimated from other sources. Estimates of the percentage of parolees that owe child support range from 28% to 32% (Herman-Stahl, Kan & McKay; McLean & Thompson). This estimate is consistent with the Utah parolee data, where 33% of parolees stated that they owed child support. In Colorado and Massachusetts, it was found that released prisoners owed an average of approximately $16,000 in arrears (Herman-Stahl, Kan & McKay, 2008; McLean & Thompson, 2007). The average monthly child support payment for those Utah parolees that owed child support was $259.48. If these payments were made for five years, the amount would be very close to $16,000. Herman-Stahl, Kan, and McKay report that the amount of child support owed by prisoners increases by approximately $5,000 on average during their time in prison. This implies that prisoners enter prison owing roughly $11,000 in child support.

If child support payments are eliminated from the Utah parolee data set, the BMA model predicts that recidivism drops by 14%. From the 221 parolees that actually returned to prison, the model predicts 31 fewer will return to prison if child support payments are eliminated. This implies that incarceration costs would drop by a total of $594,915.73. Unlike the case of restitution, the government cannot pay off all child support owed and expect reduced incarceration costs to cover the total amount of compensation. If $16,000 is a good estimate of child support owed in arrears for Utah parolees, the total amount owed by the 169 parolees who must pay child support is $2,704,000. This cost cannot be offset by the expected decrease in incarceration costs.

Policy recommendations for child support payments are not as clear-cut as they are for employment and restitution payments. The analysis in Chapter 5 confirms that child support payments increase the probability of recidivism. However, with the decrease in corrections costs from reduced recidivism being insufficient to fully compensate those owed child support, there is no easy policy recommendation that makes all parties better off. Only two recommendations appear reasonable in this case. First, because recidivism will be more likely with higher child support payments, these payments need to be reduced to a level that does not lead parolees to consider illegal means for acquiring income. While this appears to be a sound principle, determining the payment size may be difficult and it will certainly vary from one parolee to another. A second recommendation would be to prohibit the accumulation of child support payments with interest while an offender is in prison. Those offenders who owe child support payments owe an average of $11,000 in arrears when beginning their sentences. If the fact of owing $11,000 in child support played a role in leading to the commission of a crime, the problem is only exacerbated by allowing the payments and interest to accumulate while the offender earns as little as 40 cents per day in prison. Beyond these two recommendations, the issue of criminal justice policy formation with respect to child support payments is a difficult one.

In summary, three policy recommendations were made with respect to the three economic policy variables. Efforts to improve the chances of employment for released prisoners appear justified up to an annual amount of $18,172.93 for released prisoners. Any amount below this figure that successfully created employment would appear to produce a net benefit for society. Removing the burden of restitution payments from released prisoners can also benefit society. Even though restitution payments are transfers and their complete elimination would still produce net gains, a government compensation scheme could satisfy victims and still reduce overall criminal justice costs. Finally, because recidivism is influenced by the requirement to pay child support, these payments need to be reduced to a minimum level that prevents recidivism and provides some support to children. A change in the law that prevents child support payments to accumulate with interest while an offender is in prison could help reduce monthly payments significantly for released prisoners, thereby reducing the risk of recidivism.

6.2 Political impediments to policy change

While the policy recommendations above would appear to produce better outcomes for society as a whole, resistance to the implementation of the policies is more likely than not. As is the case with many other political issues, the very nature of the political system makes seemingly beneficial policy change difficult to enact. Special interest groups often lobby successfully for laws or policies that benefit the special interest group at the expense of the rest of society. Logrolling can lead to the passage of laws and policies that benefit few at the cost of many. Politicians frequently adopt the planks of their platforms based on emotional response without ever fully disclosing the implied costs of such policies. These criticisms of the political system are not new and they are generally well understood. The purpose here is not to produce a lengthy critique of political theory, but instead to identify the special interest groups and the incentive problems associated with criminal justice policy formation in particular.

The participants with possibly the greatest influence over criminal justice policy within the political system are the victims' rights and child support rights special interest groups. These groups are very well organized with representation at the federal, state, and local levels. The organization of these groups can be based on both financial and retributive motives. From the purely financial perspective, these groups have a strong incentive to push for restitution and child support payments, whereas the average taxpayer has little incentive to lobby in opposition. To use an example based on the figures from the previous section, a victim receiving restitution can expect $1,890, while the cost savings from the complete elimination of restitution implies only a $10.50 reduction per capita in the State of Utah. Thus, the general public has little incentive to reform restitution laws and victims' rights groups have a large incentive to maintain them in their favor. The same holds true for child support rights groups. It may be true that the retributive motive toward political organization among victims' rights and child support rights groups is even be more powerful. The punishment of criminal offenders is most certainly an important social function. Moreover, there are many cases where the life imprisonment of a criminal offender may be in the best interest of public safety. However, society is better served by limiting the satisfaction of the demand for retribution. The incarceration of criminal offenders on the basis of retribution when they may not pose any genuine risk to society brings personal satisfaction to some, but places a tax burden on all.

The "tough on crime" platform has been very successful at winning elections. While it may be true that all or most public officials are motivated to some degree by a desire to serve the public, they all must necessarily win elections. All members of the executive and legislative branches of government can conceivably benefit from a tough on crime platform and they uniformly dread being perceived as soft on crime. Frequently, a sensationalistic account of a heinous crime widely reported by the media will lead to irrational fears of crime stemming from an incorrect assessment of the probabilities of genuine threat as a reaction to the event. Politicians can use this fear to promise tougher crime laws in exchange for election. The irrational fear resulting from a sensationalistic crime story can lead to another equally irrational conclusion: Any cost is justified in order to place all criminals behind bars. The emotional excitement resulting from a high profile criminal event usually distracts citizens from considering the exact costs of the policies promised by an opportunistic politician. Although it is difficult to overcome emotionally-based prejudices, a possible solution to the tough on crime regime is to give a full accounting of the implied costs of such policies before elections and convert it into a dollar amount per taxpayer. Knowing the exact amount of income that will go to increased corrections spending may have a sobering effect.

The courts may also be susceptible to the criticism of placing popularity over principle with respect to criminal justice policy. In Utah, district court judges are not initially elected, but appointment on the basis of merit. After a 4-year term all judges must face a retention election. This can create an incentive for judges to develop the appearance of being tough on crime, which can be accomplished by using whatever discretion is at their disposal to impose harsher sentences. With victims being directly present through the adjudication process, the evaluation of the judge's stance on crime can be immediately reported through various victims' right groups.

The last two groups with political influence over the formation of criminal justice policy that are mentioned here are law enforcement and corrections. It is immediately apparent that these departments directly benefit in terms of higher funding from tough on crime policies. If recruitment into these agencies lags behind the demand for these public safety services, higher wages and overtime pay usually results. Incentives toward higher incarceration may exist in other forms as well. Ruback and Bergstrom (2006) note that many parole and probation departments charge fees used to cover their own costs. In Texas, 40% of probation costs are covered through fees paid by offenders (Ruback & Bergstrom). Parole and probation officers are often evaluated in terms of their ability to collect restitution, fees, and fines from parolees and probationers. This creates a strong incentive to strictly enforce these policies. Occasionally reincarcerating parolees and probationers for failing to pay provides credibility to the threat of incarceration, which in turn may serve to motivate payment by other parolees and probationers.

This discussion of the potential political difficulties involved in changing criminal justice policy is intended only for the purpose of identifying the influential groups involved within the political system and describing their incentives. A full discussion of possible solutions to these political issues is beyond the scope of this study. In general, the ideal solution to the inherent problems of representative democracy is the development of a well-informed public that understands the issues, incentives, and consequences for society as a whole and that holds all representatives accountable for their decisions. However, just as in the case of criminal justice policy, there may be strong incentives working to prevent the realization of this presumably beneficial ideal.

6.3 Directions for future recidivism research

Given the importance of economic variables in predicting recidivism and the low cost of collecting economic information relative to the potential cost savings, further research into economic factors and their possible influence on recidivism appears justified. Several economic variables that likely have a significant influence on recidivism have never been adequately studied. In addition to the further study of economic factors, several other variables within the Utah parolee data set were left unexplored as they were deemed beyond the scope of the present study. Before noting these other variables of interest in the Utah parolee data set and the other economic variables in need of future research, a few general suggestions for the improvement of modeling recidivism in Utah are provided first.

In order to obtain more accurate estimates, larger samples could be produced by extending the survey to all parolees and/or large samples of probationers. The influence of education on recidivism could likely be improved through the use of standardized test scores, lists of course taken, and grade point averages, to name a few. A better understanding of the living conditions of parolees and probationers could be obtained by collecting information on marital status, number of dependants, and age at first contact with the criminal justice system.

Greater accuracy can also be obtained through the verification of survey data by sources other than those being questioned. Information regarding employment status, wage rate, and hours worked per week are standard tax record items. Information on the amounts of restitution and child support owed could be obtained from the courts. Of course, the right to privacy of parolees and probationers should always be a primary concern, but, assuming that the rights of parolees and probationers are fully protected, alternative sources of verification would certainly improve accuracy.

A few variables in the Utah parolee data set were not analyzed within the context of this dissertation because they were not directly relevant to the main thesis. Yet, these variables may provide interesting and useful information regarding certain characteristics of parolees and the influence of such characteristics on recidivism. One such variable attempted to capture the types of hobbies, pastimes, or leisure activities in which parolees would engage. Classifying activities according to some relatively small set of categories presents several difficulties. Nevertheless, this variable could provide interesting information regarding certain types of behavior and their relationships to recidivism. A few questions in the survey were devoted to determining the reasons for why parolees were unemployed. These were not considered within the dissertation, but an analysis of these responses could be useful for determining ways to help unemployed parolees and probationers find jobs.

Regarding economic variables, the behavior of parolees and probationers will likely be influenced to an even greater extent by economic factors in the future. Ruback and Bergstrom (2006) note that economic sanctions are likely to become more prevalent in the future due to the high cost of operating corrections departments, increased sympathy for victims groups, and the need to find low-cost alternative punishments to imprisonment. These economics sanctions include restitution, fines, fees, and forfeiture. Ruback and Bergstrom list 36 types of fines and fees levied upon parolees and probationers in Pennsylvania alone. Bonczar (1997) found that approximately 85% of all probationers in the United States had to pay some type of fee as a special condition of their probation. No research could be found on the impact of fines, fees, and forfeiture on recidivism. The expectation is that these economic factors will have a strong influence on recidivism. If a greater reliance is placed on these forms of economic sanctions in the future, it will be of considerable importance to carefully estimate their impact on recidivism. Research may demonstrate that the imposition of these economic sanctions leads to incarceration costs that far outweigh the perceived benefits from the sanctions.

With criminal justice expenditures placing an ever-growing burden on state budgets, criminal justice policy needs to be founded upon a careful consideration of the behavioral consequences of economic sanctions with an eye toward reducing overall costs for society.

APPENDIX A

BMA POSTERIOR DISTRIBUTIONS

An alternative method for graphically representing model specification uncertainty and coefficient magnitude uncertainty is the plotting of the posterior distributions for the coefficients. In Table 5.2, the posterior probability that the variable belongs to the best model and the expected value of the coefficient were given numerically for each variable. The plot in Figure 5.2 gave expression to the posterior probabilities associated with each variable by showing the frequency with which the variables appear in the best models. While Figure 5.2 produces only a graphical representation of the posterior probabilities that variables are members of the best model, the posterior probability distributions for the coefficients in Figures A.1 and A.2 simultaneously express the posterior probabilities that the variables belong to the best model and the expected values of the coefficients.

A posterior coefficient distribution is produced for each of the 37 variables and the intercept term in Figures A.1 and A.2. Each distribution is centered over the expected value of coefficient and is scaled so that its maximum height is equal to the probability that the variable is a member of the best model. A spike is centered at zero in each distribution and its height expresses the probability that the variable is not a member of the best model. The sum of the maximum height of the distribution and the height of the spike is one. Together, Figures A.1 and A.2 provide a complete graphical representation of the posterior probability, expected value, and standard deviation columns in Table 5.2 for all of the variables.

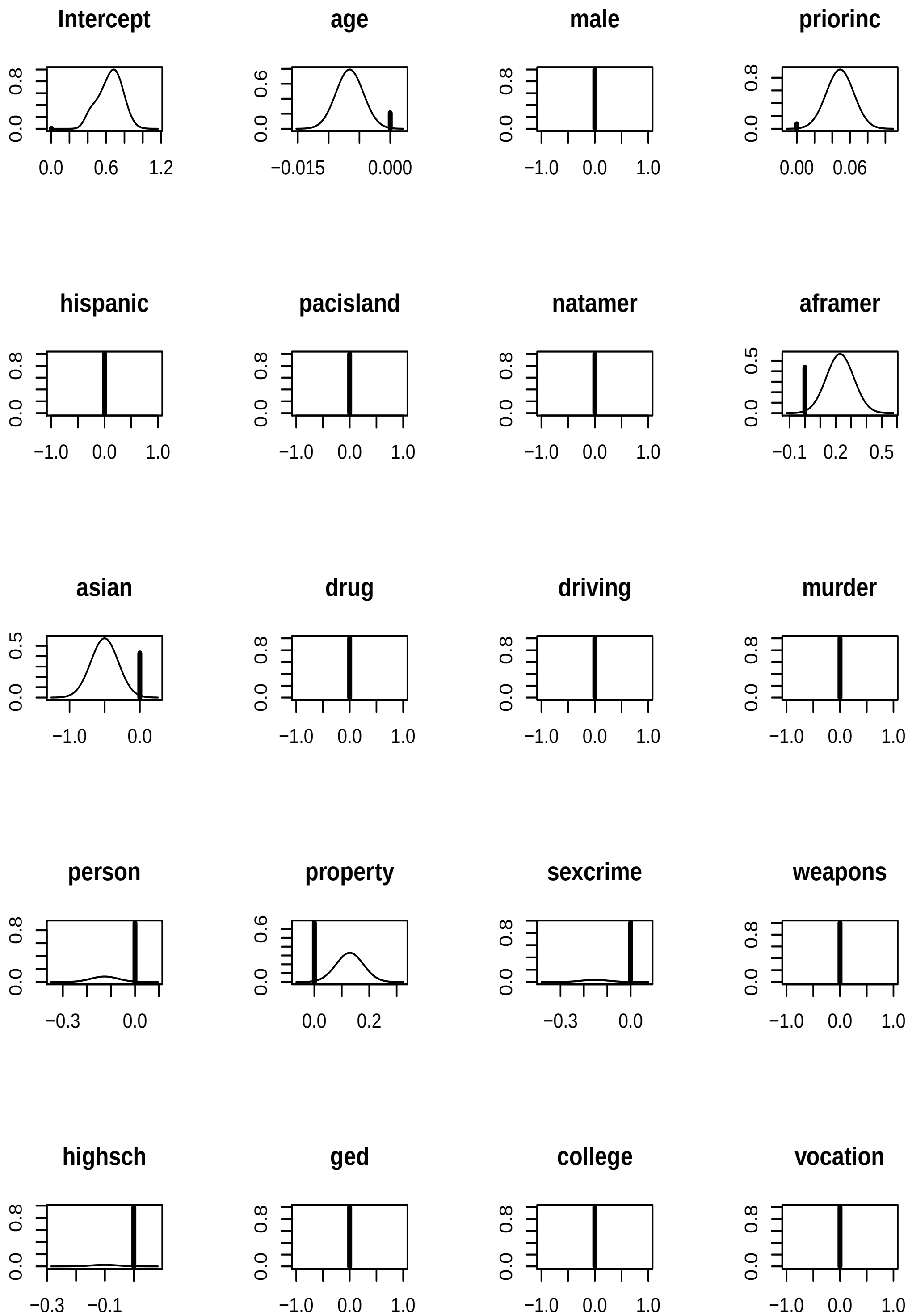

Figure A.1. Posterior Distributions for the Intercept and Variables 1-19

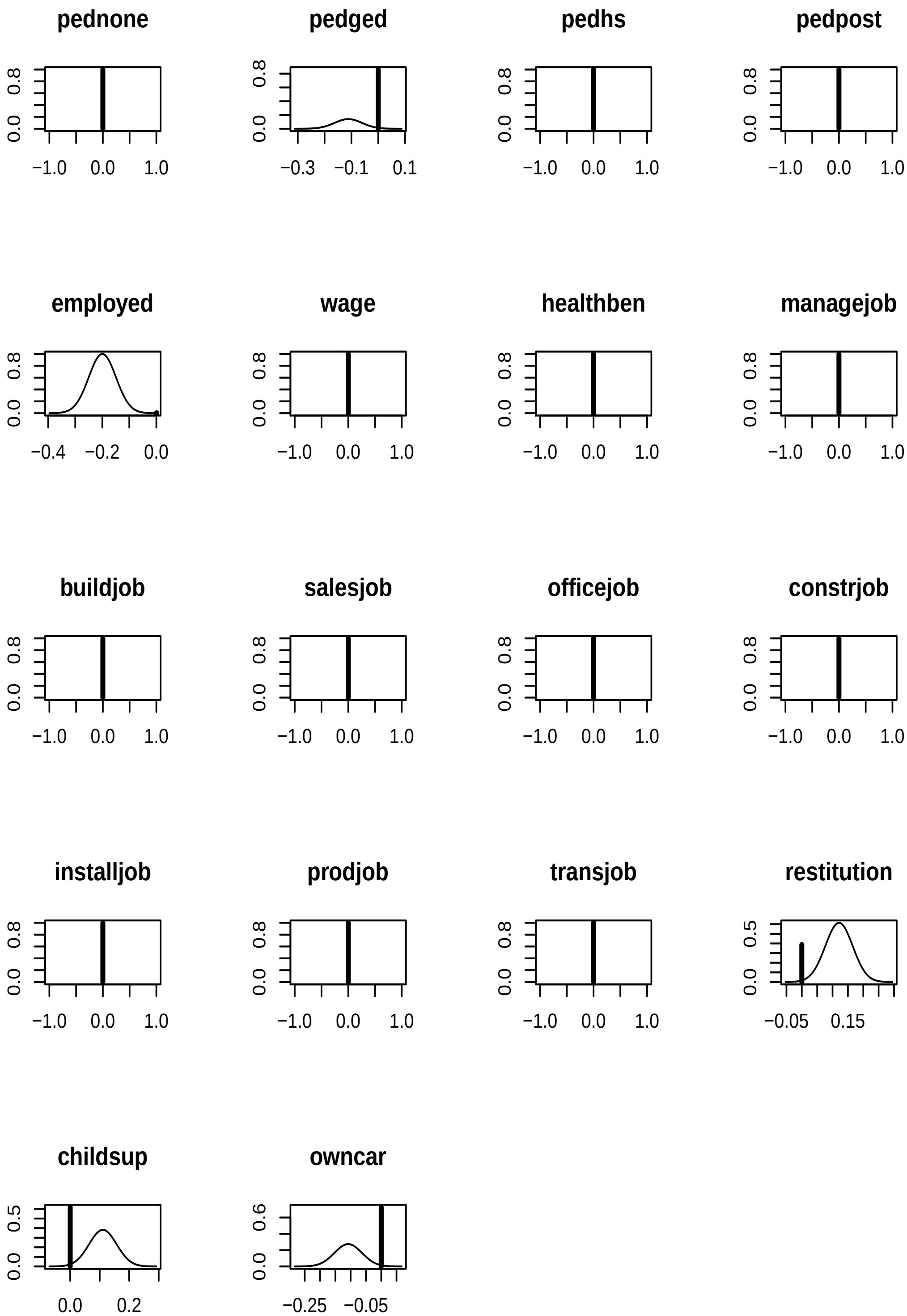

Figure A.2. Posterior Distributions for Variables 20-37

APPENDIX B

GRAPHICAL MULTIVARIATE REPRESENTATION

The graphical representation of relationships and functions involving more than a total of three variables is a difficult problem. A function of two independent variables can be represented as a surface within a three-dimensional Cartesian coordinate system, but functions of more variables cannot be represented in this same fashion. Chernoff (1973) invented a novel solution to this problem. By assigning variables to different facial characteristics, Chernoff was able to uniquely represent up to 18 variables of data as a set of human faces. Thus, an 18-dimensional data point can be represented in a single graphical image, allowing the analyst to perceive 18 characteristics simultaneously. Chernoff's graphical technique was applied to the Utah parolee data in order to investigate whether recognizable patterns exist among recidivists relative to nonrecidivists.

From the Utah parolee data set, only 20 observations were selected for graphical representation. Two subsets of the data were chosen to provide the greatest contrast between those who recidivated and those who did not. From the set of 221 parolees that returned to prison, the BMA model was used to select the 10 parolees with the highest probability of returning to prison. Likewise, from the set of 285 parolees that did not return to prison, the 10 parolees with the lowest probability of returning to prison were selected for the other subset. The faces were modeled using 15 different variables, where all of the most important predictors of recidivism were included. The graphical representation of these 20 observations is presented in Figure B.1.

Although the faces do not yield perfectly uniform patterns, several distinct similarities are detectable. The direction of the eyebrows, the width of the eyes, and the overall shape of the heads are probably the most easily recognized features

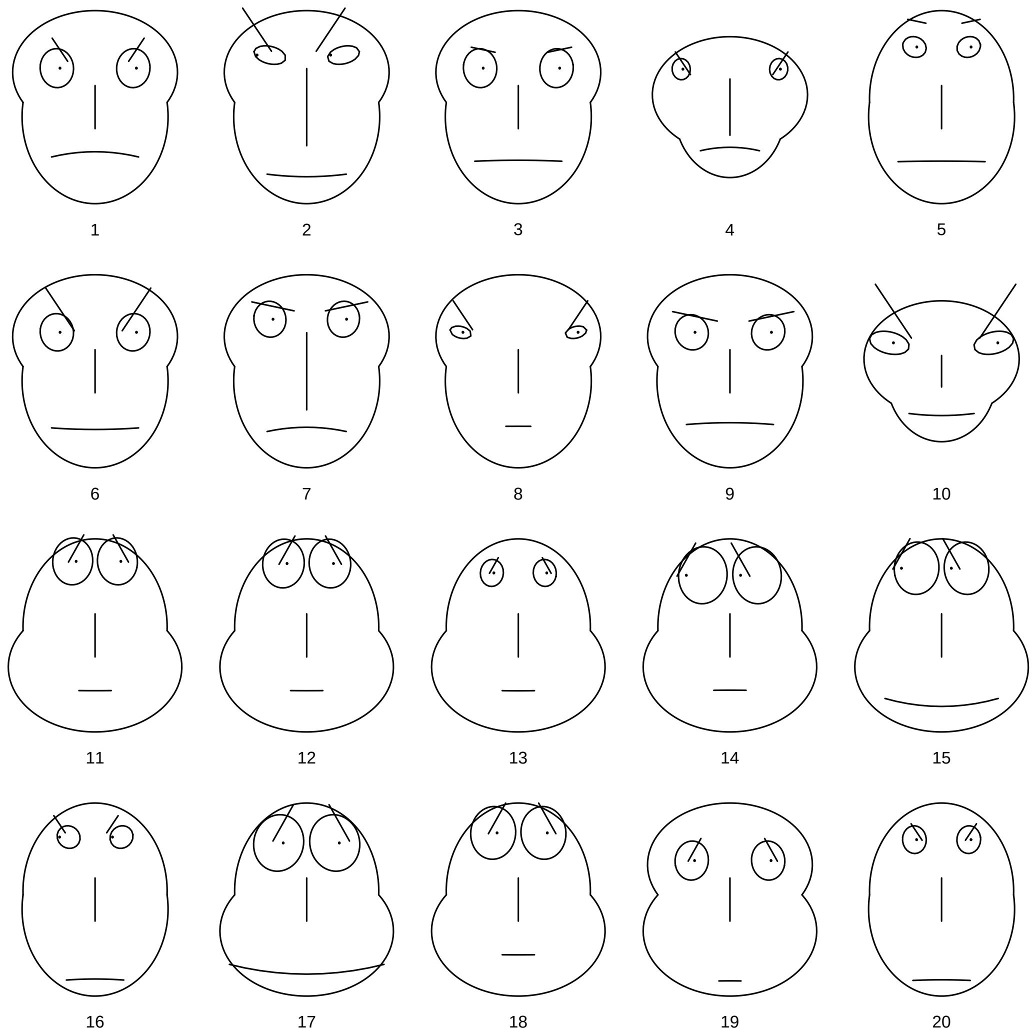

Figure B.1. Multivariate Representation of the Utah Parolee Data

that differ in a reasonably consistent manner across those who did and did not recidivate. These features were largely determined by the employment, restitution, prior incarcerations, wage, property crime, and high school diploma variables. A primary virtue of this graphical technique is its ability to reveal when certain types of individuals (i.e., recidivists and nonrecidivists) share many different characteristics. Of the 10 parolees who returned to prison, 9 were unemployed, all paid restitution, 4 had committed a property crime, 3 graduated high school, and the average number of prior incarcerations was 3.6. For the 10 parolees who did not return, 9 were employed, 2 paid restitution, none had committed a property crime, all graduated high school, and the average number of prior incarcerations was 1.4.

It is important to note that the particular type of expression conveyed by any face is meaningless. Even though recidivists and nonrecidivists appear to communicate particular emotional states (e.g., anger, sadness, worry, etc.), these states are completely arbitrary. By simply inverting the values of the variables or selecting different variables to assign to the facial characteristics, a recidivist could be made to express just about any emotional state. This graphical procedure does not produce objectively meaningful emotional expressions for a given data set. Nevertheless, the emotional expressiveness of any particular face, even when arbitrarily assigned to a random observation, does play a role in the analytical process. The purpose of using faces for this technique is to appeal to the ability of individuals to perceive subtle differences in facial expressions in order to discern patterns in the data.

REFERENCES

[1] Brendan Anstiss. Just how effective is correctional treatment at reducing re-offending? *New Zealand Journal of Psychology*, 32(2):84–91, 2003.

[2] William D. Bales and Daniel P. Mears. Inmate social ties and the transition to society: Does visitation reduce recidivism? *Journal of Research in Crime and Delinquency*, 45(3):287–321, 2008.

[3] Eric P. Baumer, Richard Wright, Kristrun Kristinsdottir, and Helgi Gunnlaugsson. Crime, shame, and recidivism: The case of Iceland. *British Journal of Criminology*, 41:40–59, 2002.

[4] Allen J. Beck and Bernard E. Shipley. *Recidivism of Prisoners Released in 1983*. U.S. Department of Justice, Bureau of Justice Statistics, Washington, D.C., 1989.

[5] Alfred Blumstein and Richard C. Larson. Problems in modeling and measuring recidivism. *Journal of Research in Crime and Delinquency*, 8(2):124–132, 1971.

[6] Thomas P. Bonczar. *Characteristics of Adults on Probation, 1995*. U.S. Department of Justice, Bureau of Justice Statistics, Washington, D.C., 1997.

[7] James Bonta, Mia Dauvergne, and Tanya Rugge. The reconviction rate of federal offenders. User report 2003-02, Solicitor General of Canada, Ottawa, Ontario, 2003.

[8] Leo Breiman, Jerome H. Friedman, Richard A. Olsen, and Charles J. Stone. *Classification and Regression Trees*. Chapman & Hall/CRC Press., Boca Raton, FL, 1984.

[9] California prison riot blamed on crowding. *msnbc*. August 10 2009.

[10] California prison spending rises, so do concerns. *msnbc*. August 23 2009.

[11] California Senate OKs ealy inmate release. *msnbc*. August 20 2009.

[12] Shannan M. Catalano. *Criminal Victimization, 2003*. U.S. Department of Justice, Bureau of Justice Statistics, Washington, D. C., 2004.

[13] Herman Chernoff. The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, 68(342):361–368, 1973.

[14] Ted Chiricos, Kelle Barrick, William Bales, and Stephanie Bontrager. The labeling of convicted felons and its consequences for recidivism. *Criminology*, 45(3):547–581, 2007.

[15] Denmark Department of Prisons and Probation. *Kriminalforsorgens Statistiks 2005*. Department of Prisons and Probation, Copenhagen, Denmark, 2006.

[16] Federal Bureau of Investigation. *Crime in the United States, 2003: Uniform Crime Reports*. U.S. Government Printing Office, Washington, D.C., 2004.

[17] Stephen Fienberg and Patricia Grambsch. An assessment of the accuracy of *The Effectiveness of Correctional Treatment*. In Lee Sechrest, Susan O. White, and Elizabeth D. Brown, editors, *The Rehabilitation of Criminal Offenders*, pages 119–147. National Academy of Sciences, Washington, D.C., 1979.

[18] George M. Furnival and Robert W. Wilson. Regression by leaps and bounds. *Technometrics*, 16(4):499–511, 1974.

[19] Paul Gendreau, Tracy Little, and Claire Goggin. A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology*, 34(4):575–607, 1996.

[20] Paige M. Harrison and Jennifer C. Karberg. *Prison and Jail Inmates at Midyear 2003*. U. S. Department of Justice, Bureau of Justice Statistics, Washington, D. C., 2004.

[21] Joe Heinz, Burt Galaway, and Joe Hudson. Restitution or parole: A follow-up study of adult offenders. *The Social Service Review*, 50(1):148–156, 1976.

[22] Mindy Herman-Stahl, Marni L. Kan, and Tasseli McKay. *Incarceration and the Family: A Review of Research and Promising Approaches for Serving Fathers and Families*. U. S. Department of Health and Human Services, Washington, D.C., 2008.

[23] Jennifer A. Hoeting, David Madigan, Adrian E. Raftery, and Chris T. Volinsky. Bayesian model averaging: A tutorial (with discussion). *Statisical Science*, 14(4):382–417, 1999.

[24] Kimmo Hypén. The released from prison in Finland 1993–2001 and the reentered. Paper presented at the Third Conference of the European Society of Criminology, Helsinki, Finland, 2003.

[25] Doris J. James. *Profile of Jail Inmates, 2002*. U. S. Department of Justice, Bureau of Justice Statistics, Washington, D. C., 2004.

[26] Jörg-Martin Jehle. Criminal justice in Germany. 4th ed., Federal Ministry of Justice, Berlin, 2005.

[27] Jörg-Martin Jehle, Wolfgang Heinz, and Peter Sutterer. *Legalbewährung nach strafrechtlichen Sanktionen: Eine kommentierte Rückfallstatistik*. Federal Ministry of Justice, Berlin, 2003.

[28] Craig Jones, Jiuzhao Hua, Neil Donnelly, Judy McHutchison, and Kyleigh Heggie. Risk of re-offending among parolees. Crime and Justice Bulletin 91, NSW Bureau of Crime Statistics and Research, Syndey, Australia, 2006.

[29] Satoshi Kanazawa. When crime rates go down, recidivism rates go up. *Psychology Today* [online], August 24 2008.

[30] Annie Kensey and Pierre V. Tournier. La récidive des sortants de prison. Cahiers de démographie pénitentiaire 15, Ministère de la Justice, Direction de l'Administration Pénitentiaire, Paris, France, 2004.

[31] Annie Kensey and Pierre V. Tournier. Sortants de prison: variabilité des risques de retour. Cahiers de démographie pénitentiaire 17, Ministère de la Justice, Direction de l'Administration Pénitentiaire, Paris, France, 2005.

[32] Il Kim, Bruce L. Benson, David W. Rasmussen, and Thomas W. Zuehlke. An economic analysis of recidivism among drug offenders. *Southern Economic Journal*, 60(1):169–183, 1993.

[33] Patrick A. Langan and David J. Levin. *Recidivism of Prisoners Released in 1994*. U.S. Department of Justice, Bureau of Justice Statistics, Washington, D.C., 2002.

[34] Jeff Latimer, Craig Dowden, and Danielle Muise. The effectiveness of restorative justice practices: A meta-analysis. *The Prison Journal*, 85(2):127–144, 2005.

[35] Michael D. Maltz. *Recidivism*. Academic Press. (Original work published 1984), Orlando, FL, 2001.

[36] Robert Martinson. What works? questions and answers about prison reform. *The Public Interest*, 35:22–54, 1974.

[37] Rachel L. McLean and Michael D. Thompson. *Repaying Debts*. Council of State Governments Justice Center, New York, 2007.

[38] Ministry of Justice, Research and Documentation Centre. European sourcebook of crime and criminal justice statistics - 2003. Research and policy series, no. 241, Author, The Hague, 2006.

[39] Terrie E. Moffitt. Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4):674–701, 1993.

[40] Arul Nadesu. *Reconviction patterns of released prisoners: A 36-months follow-up analysis*. Department of Corrections, Wellington, New Zealand, 2007.

[41] Arul Nadesu. *Reconviction patterns of released prisoners: A 48-months follow-up analysis*. Department of Corrections, Wellington, New Zealand, 2008.

[42] Arul Nadesu. *Reconviction patterns of released prisoners: A 60-months follow-up analysis*. Department of Corrections, Wellington, New Zealand, 2009.

[43] National Council for Crime Prevention. *Kriminalstatistik 2007*. National Council for Crime Prevention, Stockholm, Sweden, 2008.

[44] New Zealand Department of Corrections. *Census of Prison Inmates and Home Detainees, 2003*. Department of Corrections, Wellington, New Zealand, 2004.

[45] New Zealand Department of Corrections. *Annual Report for 2004/05*. Department of Corrections, Wellington, New Zealand, 2005.

[46] New Zealand Police. *New Zealand Crime Statistics 2003*. New Zealand Police, Office of the Police Commissioner, Wellington, New Zealand, 2004.

[47] Ian O'Donnell, Eric P. Baumer, and Nicola Hughes. Recidivism in the Republic of Ireland. *Criminology and Criminal Justice*, 8(2):123–146, 2008.

[48] Jason Payne. Recidivism in Australia: Findings and future research. Research and Public Policy Series 80, Australian Institute of Criminology, Canberra ACT, Australia, 2007.

[49] Pew Center on the States. *One in 100: Behind Bars in America 2008*. Author, Washington, D.C., 2008.

[50] David A. Pritchard. Stable predictors of recidivism: A summary. *Criminology*, 17(1):15–21, 1979.

[51] Adrian E. Raftery, David Madigan, and Jennifer A. Hoeting. Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92(437):179–191, 1997.

[52] R. Barry Ruback and Mark H. Bergstrom. Economic sanctions in criminal justice: Purposes, effects, and implications. *Criminal Justice and Behavior*, 33(2):242–273, 2006.

[53] D. Ruddy and A. Brown. Adult reconviction in Northern Ireland 2004. Research and statistical bulletin 6/2008, Northern Ireland Office, Criminal Justice Directorate, Belfast, 2008.

[54] William J. Sabol and Heather Couture. *Prison Inmates at Midyear, 2007*. U.S. Department of Justice, Bureau of Justice Statistics, Washington, D. C., 2008.

[55] Anne L. Schneider. Restitution and recidivism rates of juvenile offenders: Results from four experimental studies. *Criminology*, 24(3):533–552, 1986.

[56] Scottish Executive. Reconvictions of offenders discharged from custody or given non-custodial sentences in 1999, Scotland. Statistical Bulletin CrJ/2005/7, Scottish Executive Justice Department, Edinburgh, 2005.

[57] Lee Sechrest, Susan O. White, and Elizabeth D. Brown, editors. *The Rehabilitation of Criminal Offenders: Problems and Prosects*. National Academy of Sciences, Washington, D.C., 1979.

[58] Keith Spicer and Alison Glicksman. Adult reconviction: Results from the 2001 cohort. Home office online report 59/04, Home Office, London, 2004.

[59] State of Utah Governor's Office of Budget and Planning. *Governor's Summary of Legislative Action for Fiscal Year 1988-89*. Author, Salt Lake City, UT, 1988.

[60] State of Utah Governor's Office of Budget and Planning. *Budget Summary Fiscal Year 2010*. Author, Salt Lake City, UT, 2009.

[61] Statistics Austria. *Gerichtliche Kriminalstatistik 2007*. Statistik Austria, Bundesanstalt Statistik Österreich, Vienna, 2008.

[62] Statistics Norway. Recidivism in the five-year period 2001–2005 among persons charged in 2000. Statistical tables 21–24, Statistisk sentralbyrå, Oslo, 2007.

[63] Renate Storz. *Rückfall nach Strafvollzug: Rückfallraten*. Bundesamt für Statistik, Bern, 1997.

[64] Hung-En Sung. Rehabilitating felony drug offenders through job development: A look into a prosecutor-led diversion program. *The Prison Journal*, 81(2):271–286, 2001.

[65] United Nations. *Ninth United Nations Survey of Crime Trends and Operations of Criminal Justice Systems, covering the period 2003–2004*. United Nations Office On Drugs and Crime, Vienna: Austria, 2007.

[66] U.S. Census Bureau. *Statistical Abstract of the United States*. U.S. Government Printing Office, Washington, D.C., 101 edition, 1980.

[67] U.S. Census Bureau. *2007 Census of Government Finance*. Author, Washington, D.C., 2009.

[68] U.S. Department of Justice. New directions from the field: Victims' rights and services for the 21st century - restitution. Ncj 172825, U.S. Department of Justice, Office for Victims of Crime, Washington,D.C., 1998.

[69] B. S. J. Wartna and L. T. J. Nijssen. National studies on recidivism. Fact sheet 2006-11, WODC, The Hague, 2006.

[70] B. S. J. Wartna, N. Tollenaar, and A. A. M. Essers. Door na de gevangenis: Een cijfermatig overzicht van de strafrechtelijke recidive onder exgedetineerden. Research and policy series, no. 228, WODC, The Hague, 2005.

[71] Kristen M. Zgoba, Sabrina Haugebrook, and Krista Jenkins. The influence of GED obtainment on inmate release outcome. *Criminal Justice and Behavior*, 35(3):375–387, 2008.