ABSTRACT

Title of Dissertation: COST MATTERS: APPLICATION AND ADVANCEMENT OF ECONOMIC METHODS TO INFORM POLICY CHOICE IN CRIMINOLOGY

David M. Bierie, Doctor of Philosophy, 2007

Dissertation directed by: Professor Doris L. MacKenzie
Department of Criminology and Criminal Justice

This project develops and applies a cost-benefit analytic framework to evaluate a specific policy option facing the state of Maryland: To operate an early release program for adult inmates within a therapeutic boot camp facility, or a traditional prison that also emphasizes treatment. Drawing on a randomized experiment in which inmates were assigned to serve six-month terms at one of the two facilities, the study focuses on costs of administering programs and costs of recidivism during the observed 1 – 4 years after release. The data demonstrate the boot camp costs less to operate than the comparison site, and also generates significant reductions in the 'harm' incurred through recidivism. Thus, the data suggest the boot camp option generates a greater net social value for the state and community. These findings are robust to variation in assumptions and computational techniques, both standard to the cost-benefit literature as well as new approaches introduced in this dissertation.

COST MATTERS:
APPLICATION AND ADVANCEMENT OF ECONOMIC METHODS TO INFORM POLICY CHOICE IN CRIMINOLOGY

by

David M. Bierie

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Professor Doris L. MacKenzie, Chair
Professor Denise Gottfredson
Professor Gary Gottfredson
Professor David Weisburd
Professor Laura Dugan

ACKNOWLEDGMENTS

I'm grateful to a large number of people who have contributed to this work and my graduate experience. Doris MacKenzie has been the ideal advisor. She treats her students with care and patience, but also demands the very best from them. Her guidance has a wonderful impact on her students, both within and outside of the academic arena. My dissertation committee has created a valuable experience for me, helping me to fine tune the content and delivery of this research. Likewise, John Roman and Mark Cohen each dedicated a great amount of their time, energy, and good cheer to this dissertation. I am grateful for having had a chance to work with each of them.

I would also like to acknowledge professors John Laub and Doug Smith. Two more inspiring scholars would be hard to find. I am grateful for their time, guidance, and friendship over the past six years of graduate school. Likewise, I had the good fortune to have met, worked with, and traveled through graduate school with an amazing group of friends. Of course this includes the twins (Bob and Gary), my cohort (Jill and Lee), as well as Summer, Derrick, Nancy, Carole and many other great friends. And, of course, little of the work that emerges from LeFrak would have been possible without the assistance of Welf Dixon.

Along the same lines, I would like to thank my family for their support of my decade long pursuit of university life. Their support has been enduring and appreciated. This is true of my mother Arlene, as well as my grandparents, Murray and Miriam, and especially my brother, Brian. It is also true of my newer 'Ferstl' Family. Most important, I would like to thank Karen. Her love, support, patience and good humor have been invaluable. Particularly important has been her insistence that there is more to life than data.

Finally I would like to acknowledge that this project was supported by 2003-DB-BX-0004, awarded by the US Department of Justice. The Assistant Attorney General, Office of Justice Programs, Coordinates the activities of the program offices and bureaus. Points of view or opinions contained within this document are those of the author and do not necessarily represent the official position or policies of the US Department of Justice. Funding was provided by the State of Maryland, Governor's Office of Crime Control and Prevention, BYRN-2003-1063. All points of view in this document are those of the author and do not necessarily represent the position of any State agency.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

The state of Maryland operates a correctional boot camp for adults; a facility focused on delivering intensive therapy through a highly structured environment. This program allows offenders who are first time inmates and have non-violent convictions to reduce their prison terms (ranging from 1-5 years) to a six-month sentence in exchange for serving those six months successfully at the Toulson Correctional Boot Camp. This is referred to as the "MAP" contract program.

Currently, the state of Maryland is considering a policy change in which the MAP program would be administered at a traditional prison facility, rather than the boot camp. The MAP inmates would still earn early release by participating in therapeutic programs, and the *programming* would be similar to that now offered at the boot camp (i.e. substance abuse therapy, education, and life skills training). However, the MAP program would be administered at a traditional prison and therefore exclude the boot camp elements (atmosphere, rules, and management style).

Ultimately, the potential change in policy is motivated by the emergent national literature suggesting that boot camps do not reduce recidivism, relative to traditional prison. The state is still interested in the 'secondary' benefit of cost reduction associated with early release for inmates who participate in boot camp, which save money even when recidivism is equivalent because of reduced time spent detained. However, they recognize that such an early release program does not have to operate in tandem with a boot camp; it could operate within a traditional prison instead (i.e. if the traditional prison was cheaper and recidivism was similar). Therefore, the motivation for discontinuing the

boot camp program is tied to the assumption that (a) the MAP "early release" program would be cheaper to run at a traditional prison than the boot camp, and (b) that the recidivism rates would not be significantly different if the change was made.

The state of Maryland is delaying their decision on the potential change until a cost-benefit analysis of the two options is completed (the current dissertation). The key goal of the project, then, is to identify:

A. The difference in operating costs for running the MAP program at the boot camp versus the traditional prison,

B. The difference in post-program costs associated with the two policy options (i.e. financial difference derived from any change in recidivism), and,

C. Whether the benefits of the change outweigh the costs

To address these questions, the present study utilizes an experiment in which inmates who were eligible for the boot camp MAP program were randomized to serve their term at the boot camp or a traditional prison. At each site, the facility administered the MAP program to the best of their ability (i.e. provided education, substance abuse therapy and life skills training). Recidivism data was collected after release, with subjects at risk an average of just over 800 days since release. Although time at risk varied for individuals (resulting from a staggered entrance into the study, but s single follow up cut-off date) the average time at risk was roughly identical for the two groups. The current research focuses on 'new arrests' in order to quantify differences in costs of recidivism. Combining this information with detailed facility budget data, this paper estimates the costs and benefits of the two policy options: run the MAP program at the boot camp, or run the MAP program at a traditional prison.

This introduction proceeds by reviewing key issues in the boot camp literature, as well as debate regarding their utility. Because this debate is tightly linked to issues of recidivism impact, this section focuses on limitations in the recidivism-literature which tend to preclude our ability to draw strong conclusions. For example, there has been few randomized experiments, limited statistical tools applied and negligible examination of the quality of recidivism (i.e. cost-benefit methodologies). Developing this final point, the section concludes by discussing "cost-benefit" methods within criminology. In particular, it focuses on emergent methodological issues pertaining to valuation of crime and computation of monetary costs to different groups (i.e. primary and secondary victims).

Although this dissertation does not present a statistical comparison of rates or hazards of recidivism, literature reviewed refers to these types of analysis in order to paint a backdrop of what is known and the current state of the debate. It illustrates several limitations in the literature, one of which is the lack of cost benefit methodologies; the point of departure for this dissertation. However, it should be clear at the outset that this dissertation can not settle the debate on boot camps, or even contribute more than one data point in a potential meta-analysis. Any differences observed between the two policy options in this paper may be due to differences in one being a boot camp, and one being a traditional prison, or they may be derived from myriad other factors that are also different between groups (such as facility size, crowding, etc.). That being said, the fact remains that one is a boot camp, with all the managerial and social baggage that implies (such as more control over inmates, less amenities, and less crowding), and the other is a traditional prison, with the characteristics, social milieu etc. that this terms

implies. In that sense, one can not divorce this research from the context of facility type and the broader boot camp literature.

Boot Camps: Key Issues and Debate

The term "correctional boot camp" refers to a prison management system in which a highly structured environment is created for inmates; an environment distinguished by its paramilitary interaction style, relatively intense level of structure for inmate schedules, and restricted privileges given to inmates (i.e. free time, televisions, personal clothing, etc.). Modeled after military boot camps, the early forms of this intervention emphasized physical discipline and drill-and-ceremony. These early models were grounded in deterrence theories and "get tough" doctrines of correctional management. More recently, the boot camp model has evolved into a treatment oriented paradigm in which regimentation is seen as a way to create an environment conducive to program implementation and safe facility management, rather than an end in and of itself (MacKenzie and Armstrong, 2004).[1]

Boot camps were first introduced into the US correctional system in Oklahoma and Georgia in 1983 (MacKenzie and Armstrong 2004). Since those first two camps

[1] Indeed, there is an additional difference between the early and recent versions of boot camps; a difference which is intriguing because it is a difference in the *meaning* of actions by staff more so than a change in staff *actions*. In the early approaches, the actual 'pain' of aggressive guards and the 'pain' of physically demanding exercise or schedules was thought a key element of the correctional process; a process by which offenders would be deterred from future crime because of an increase in their perceived severity of the prison experience. Recent approaches to the camps do not wholly divorce themselves from this possibility. But it is no longer a central or dominant path by which drill instructors believe their behavior matters. Rather, drill instructors often view their impact as one of cognitive change and habit-formation. They see the structure and discipline as a way to show inmates how to live a structured life, and that they are capable of living a structured life. They see their job as teaching inmates positive thinking patterns, positive attitudes, and cognitive 'tricks' to use when facing adversity. Through the use of these tools and skills, they hope the inmates who choose to "go straight" will have practice and experience at living a crime-free lifestyle (i.e. know how to get up on time for a job, how to use professional manners and display a positive attitude to employers, how to control themselves when upset, etc.). In short, the drill instructors of old and the drill instructors of new are all tough, confrontational, and yell a lot. But the reason they yell, the specific content of what is yelled, and the meaning that is behind the confrontation has changed in ways that inmates are sure to notice – creating a substantively different experience for staff and inmates alike.

were opened, the intervention has grown exponentially. By 1994, 30 state-level camps were operating and by 1995, 52 state-level boot camps were in operation. Starting in 1990, the correctional option was expanded to juveniles, growing to at least 56 operating juvenile boot camps by 2000. These estimates do not include the more numerous private and county run camps that also began opening in the same time period, making the above count conservative (Armstrong, 2004).

The quick emergence of boot camps derived from the timing of their introduction as well as the sustainability of the option to diverse political interests. Conservatives were quick to endorse them as "tough on crime;" reducing amenities to inmates and making the experience of incarceration more harsh. Liberals were quick to endorse them as avenues to increased safety, increased access to therapy, and a tool to reduce growing expenditures on prisons derived from the newly emerging war on drugs (MacKenzie, 2001). With this context, the camps found fast and strong support across the U.S.

However, the intervention has not been without controversy. Opponents of the intervention argued that the camp atmosphere fostered a staff culture which was too aggressive and potentially abusive–creating an environment antithetical to therapeutic delivery. For example, some worry that the aggression endemic in staff-inmate interactions at camps may lead to deligitimating views of staff and animosity towards authority (see Sherman, 1993). Likewise, opponents argued that the atmosphere may generate 'harmful stress,' leading to maladjustment, depression, anxiety, and acting out. For example, Gover (2003) found that juveniles with histories of child abuse fared worse over time with respect to psychological functioning in boot camps relative to traditional facilities. Her analysis implied the experience of aggressive interaction with authority

figures could lead to deleterious outcomes for inmates, at least for this specific subpopulation. Summarizing this groups views, Morash and Rucker (1990) ask why policymakers expect an intervention designed to turn normal people into killers (i.e. military boot camps) to turn offenders into complacent citizens.

Proponents of the camps have not been silent in this debate. First, they have been quick to point out the term "boot camp" has been applied to dramatically different types of interventions. This makes bold and all encompassing conclusions of "they don't work" or "they're abusive" inappropriate. Second, although proponents also agree that abuse by staff is harmful, they do not see abuse as inherent to the boot camp philosophy. Abuse by staff can happen in any prison management system–and the evidence does not imply it is any more prevalent in boot camps than in other prison systems. For example, MacKenzie et al. (2001) examined approximately 50 juvenile institutions across the U.S., half of which were boot camps. They found that "fear of staff" was no different across facility types. In addition, they found that juveniles at boot camps viewed staff more favorably than juveniles at other types of institutions.

A third difference is that proponents are more likely to see some forms of stress as helpful, rather than assuming all stress is harmful; particularly when speaking to criminological change. Proponents argue that specific forms of stress can be healthy, providing motivation for change and fostering susceptibility to intervention (MacKenzie and Armstrong, 2004). Indeed, this argument is consistent with several theoretical traditions within criminology. This includes correctional cognitive-behavioral paradigms, in which consistent accountability, feedback and conditioning is emphasized in order to motivate enduring change. Likewise, the idea that stress can lead to change in

cognitions is consistent with rational choice theories of offending, in which stress associated with sanctions may lead to 'updating' perceptions of the costs of crime (see Pagorsky et al., 2004). Finally, the notion that stress can foster change is also consistent with General Strain Theory's conceptions of pathways by which thinking patterns, values, and beliefs change to alleviate emotional duress (see Agnew, 1992; 2001). Proponents, then, have argued that the 'stressful' nature of boot camps can be integral to fostering pro-social change. In contrast, opponents have been (a) skeptical of the link between this form of stress and change, and (b) worry that the aggressive interaction can be a slippery slope which may too easily lead to injustice and abuse.

Proponents disagree with this later assertion in particular, arguing that camps are at least as safe as traditional prisons, if not more so. For example, the literature implies that camps generate lower rates of exploitation and greater feelings of safety among inmates (see MacKenzie et al. 2001). Proponents argue this is a valuable goal in and of itself. However, they also argue that greater control likely has *indirect* benefits beyond mere safety. That is, the added structure, regimentation, physical exertion, and the esprit de corps among staff may play an important role in creating the foundation on which to deliver effective programming (see Mitchell et al. 2001). It may allow for better managed classrooms, more motivated participation in classrooms, and better delivery by staff.

More specifically, the literature suggests environments which have higher rates of violence, exploitation, and fear can foster important cognitive and social changes among inmates; such as a anti-authority views or hyper-masculine persona designed to ward off predators (see DiIulio, 1987; Anderson, 1999; and Goffman's, 1967 discussion of face-

work). The 'fronting' and 'preemptive violence' characteristic of prisons with higher rates of fear can be a potent impediment to effective service delivery. Not only are resources shifted away from providing materials for treatment (i.e. shifted to hiring more security personnel rather than more vocational teachers), but the mere presence of a fear-based-culture reduces the ability of classrooms to be managed, as well as willingness of inmates to participate in treatment. That is, sincere participation in programming can appear "weak" in the eyes of predators, opening individuals to exploitation. For these reasons, the added structure, regimentation, and the associated feelings of greater safety imply that the camps may foster more effective platforms for service delivery.

However, it is not difficult to see that some (particularly policymakers) feel frustrated by the debate; seeing it as semantic or academic. There are some who would argue that the 'theoretical consistency' of the boot camp model with General Strain Theory, or the 'plausibility' that regimentation fosters more effective frameworks for program delivery is beside the point. Rather, they want to know an answer to one key question: Do they work?

Do Boot Camps Work?

There are myriad goals for boot camps. They are implemented to increase safety, foster prosocial change in psychological functioning, and reduce prison crowding. Every goal is important, and literature has accumulated on each of them (see Mackenzie and Armstrong, 2004 for a review). But there is one goal which stands out; a key question as far as policymakers are concerned: Do they reduce recidivism?

The answer to that question must, of course, be qualified. The answer is a function of (a) the quality of research conducted, and (b) the quality of the program evaluated (i.e. the subpopulation of inmates referred to and the characteristics of camps examined).

First, the camps differ across geographic space, and certainly they have changed over time. Specifically, they vary in the populations targeted (i.e. high risk versus low risk; men versus women, children versus adults), and in the quality of facilities and staff (i.e. staff to inmate ratios, training and experience of staff, whether housed in a separate facility or a subunit within a traditional prison). Likewise, they vary in program design (i.e. emphasis on 'drill' versus 'therapy') and fidelity of implementation. Evaluations which examine a single program or type of programs should not be used to generalize to substantively different programs or populations.

Second, most evaluations use quasi-experimental designs that suffer from selection bias; a problem which can be particularly difficult to correct with the small samples endemic in this literature and lack of creative or rigorous methods to model the selection process. Others suffer from inadequately addressed attrition problems. Still others have inappropriate follow up periods, limited outcome measures, or statistical procedures which are ill-fitted for specific data sets. Although the literature tends towards a finding of null-effects, the conclusions are tentative; awaiting clear methods and strong statistical tools to allow more definitive statements. Keeping these two caveats in mind, the recidivism literature is reviewed below.

With regard to recidivism, the literature implies that boot camps do not, on average, work. Peters (1996) examined three juvenile boot camps implemented by NIJ as

part of a multi-site demonstration project (see also Thomas and Peters, 1996). The studies utilized a randomized design in which juveniles were sent either to the boot camp or a traditional prison. They found that recidivism was no different in two sites, and increased among boot campers in the third site. Some, however, have been hesitant to endorse these findings as conclusive – noting that these boot camps (a) were in the first stages of implementation and suffered in fidelity for this reason, and (b) the camps emphasized discipline and punishment rather than therapy (Bourque et al., 1996). In fact, one site had to be shut down prematurely due to abuse of children by staff, an event reinforcing the suggestion that that these may not have been reasonable facilities to use as a representative of "boot camps" in general.

In the only additional randomized experiment, Bottcher and Ezell (2005) compared the time to first arrest between juveniles assigned to boot camp versus traditional facilities in California. The boot camp was also a newly implemented site, evaluated at the same time as the three NIJ funded sites. Here, the authors found no difference in recidivism. In sum, the extant literature contains only these two separate studies using randomized designs, both focusing on juveniles and newly implemented boot camps operating in the early 1990s.

Turning to adult boot camps, MacKenzie and Brame (1995) examined recidivism differences across eight quasi-experimental boot camp evaluations. They found that, on average, there was no reduction in recidivism as compared to traditional incarceration. However, they also noted that recidivism *was* lower among those facilities which dedicated a significant amount of their time to treatment rather than drill-and-ceremony.

This research lent credence to arguments made by boot camp proponents; they may be a viable option if they are therapeutic-based.

The most comprehensive research examining recidivism and boot camps to date was a meta-analysis completed by Wilson, MacKenzie, and Kider (2001). These authors explicitly accounted for differences in (a) research design quality and (b) characteristics of boot camps (such as therapeutic intensity) when reviewing all known studies of boot camps. Identifying 32 unique research studies (and 43 unique samples within them), they found the overwhelming story in the literature was one of null effects. Of the thirty-two studies, five found boot campers recidivated less than controls, four found they recidivated more, and the remainder showed no difference.

The literature implies that boot camps do not, on average, reduce recidivism. But this finding is based on a literature notorious for limited rigor in research methodology (i.e. research design, as well as tools used to analyze data). Only two randomized studies exist. The first examined juvenile boot camps which were poorly implemented, focused on physical exertion rather than therapeutic delivery, and were so poorly managed that one had to be shut down during the study due to abuse by staff. The second also examined a juvenile camp and emphasized drill and ceremony rather than therapy. The meta-analysis techniques are a strong tool from which to account for differences in facilities, samples, and research quality when reviewing the literature. The technique helps us makes sense of what the literature says and how confident we can be in conclusions. What the technique can *not* do is take the place of high quality research studies from which to draw those conclusions (i.e. analyzing 30 poor studies will not tell us what good research would have found). The present study, then, fills an important gap

in the field with respect to high quality experiments to determine whether a well run, well implemented, and therapeutically intense boot camp can reduce crime relative to a traditional prison.

The literature is limited not only by a lack of randomized designs, but also by deficiencies in analytic procedures. Two statistical gaps permeate the literature and are worth noting here. First, there has been a dearth of studies utilizing survival analysis to test for differences in recidivism (see Botcher and Ezell 2005 for an important exception). In contrast, most analysts use logistic regression to compare proportions of recidivism in two or more groups, a tool which does not account for the timing of events. It is possible (and even plausible) to see two programs which generate similar proportions of inmates who recidivate, in the long run, yet generate substantive differences in the time to that failure rate. This scenario would imply there were real and significant differences in the programs such that one generated less crime – but that insight would be lost because of the failure to use a statistical tool which can identify it.

A second important gap in the boot camp literature has been the lack of cost-benefit analyses within program evaluations. Aos et al. (2001) provide an important exception. They completed a meta-analysis of 11 boot camp evaluations, computed an effect size for the camps, and attached estimates of criminal justice savings to that effect for a hypothetical deployment of the camps across the state of Washington. As the analysis was hypothetical, they had to make bold assumptions regarding the generalizability of the recidivism effect (i.e. what rate would be observed, the distribution of offense types among recidivists, the rate of conviction and incarceration derived from those hypothetical crime types, etc.). However, other than this "hypothetical" evaluation,

program evaluations of actual boot camps are nonexistent. This is problematic because cost-benefit techniques allow more finely grained measures of the quality of recidivism than logistic regression or survival analysis alone.

Cost-benefit techniques are not just tools for accountants; they represent critical tools in program evaluation. The methods can generate more accurate and thorough program evaluations than simple analysis of event differences because they offers a strong framework for computing differences in the *quality* of recidivism (see Welsh and Farrington, 2001; Cohen 2001). Not all recidivism events are the same. Yet the literature rarely has a good means for dealing with this; for "weighting" crime events. At best, studies may look at violent versus non violent recidivism, or reincarcerations versus more minor penalties. As illustrated by Cohen (1988), cost-benefit analysis is a reasonable tool for calculating crime severity; offering far more detail and precision in the way "weighting" values are computed and attached to recidivism events. Turning to a more thorough discussion of cost-benefit analysis, the following section describes the key aspects of the method and outlines its contributions to criminological program evaluation.

Cost-Benefit Criminology: Introduction & General Framework

Criminological cost-benefit analysis is a systematic process by which monetary values are attached to the cost and return for investments in criminal justice interventions. The method is explicitly designed to inform policymakers regarding the allocation of scarce tax dollars. The approach treats criminal justice decision making as analogous to "investment" decisions; a process in which information on returns is necessary to inform decisions – decisions which are not always straightforward:

> Like any investor considering modifications to his or her investment portfolio, criminal justice policymakers need bottom-line information in order to understand and compare the resource implications of their options and decisions. For example, is it worthwhile to spend $5,000 per juvenile offender on a treatment program if recidivism is reduced by 50 percent? Is it worth it if recidivism is reduced by only 5 percent? How about 1 percent? (Aos et al. 2001, p. 150)[2]

In the realm of cost-benefit "program evaluations," there is a general framework which criminological research follows. Analysts begin by (a) computing a preliminary a statistical comparison of events (a traditional statistical comparison). Next, they turn to (b) the actual cost-benefit analysis, a process consisting of three steps.[3]

In the preliminary analysis, a basic program evaluation is conducted. In most cases, this involves a comparison of treatment and control groups in terms of key dependant variables that may later be 'valuated.' For example, this analysis may focus on determining whether an intervention generated a significant reduction in recidivism as compared to a control group. Most authors suggest the cost-benefit study itself is only justified if the program evaluation passes this first test (suggesting it makes little sense to compare the difference between the two programs in terms of the monetary value of these

[2] No doubt, there are some who would argue any reduction in crime – even a tiny one – is worth large expenditures. These people would put almost an infinite weight on victim pain and suffering, which likely drives the interpretation. However, such a position often fails to recognize that decisions are not presented in a social vacuum; resources are limited and a deployment implies a choice of *where* money should be spent. Would we choose the program with 1% reduction if that means money can not be sent to an alternate prevention program, or to other social programs (i.e. schools, homeless shelters, etc.)? Decisions surrounding interventions can not be divorced from the social world in which tax-dollars are scarce. Almost every investment in one program implies reductions in another, which means that the question is not just "what is the value of saving a victim?" Rather, it is more accurately, "what is the best way to save victims?" see Cohen 2001, as well as Boardman et al. 2001 for further discussion of 'opportunity costs.'

[3] Specific texts may outline more or fewer steps in their outline of cost-benefit analysis. Regardless of how they are broken down, they result in series of steps which are accurately captured here in the four steps listed (see Cohen 2005; Cohen 2001; Boardman et al. 2001).

variables when no difference exists).[4] In fact, the published cost-benefit evaluations almost exclusively contain examples of treatment programs that "work" – they rarely contain examples of programs which do not work or are iatrogenic (Welsh and Farrington, 2001). Again, this is because analysts usually will not complete cost-benefit studies under these conditions (the point is considered moot). However, if there are any differences identified between the treatment program and a control group among key variables (such as recidivism), then the three steps of the "cost-benefit framework" are completed.

The inclusion of this "preliminary step" within cost-benefit studies is an unfortunate trend in the field. It fails to capitalize on a key proposition of cost-benefit research: types of events are weighted differently and this can generate a substantive difference even when rates of events are the same. Two programs could have equal recidivism rates, but the severity could diverge – implying important policy information and serve as a solid foundation on which to build a cost-benefit analysis. Likewise, the actual post-program costs could be identical (same recidivism costs), but the in-program costs could be very different. In this case, a lack of cost-benefit research would again lead to a failure to produce useful policy information. For these reasons, the dissertation endorses an approach to cost-benefit methods that focuses on the final three steps alone:

[4] This step can be quite complex. Each dependant variable that may be used in later steps to represent costs (i.e. differences in rearrest, reconviction, and reincarceration – or differences in other variables monetary estimates will be attached to) could be compared using an appropriate statistical method to determine whether any differences exist. This process of statistical comparison of events would continue until at least one is found to be different; thus justifying the exploration of costs and benefits. Because of the different forms these potential variables may take (i.e. dichotomous recidivism events, differences in wages, etc.) this step of the framework can represent a rigorous undertaking. However, most published research does not undertake more than a cursory glance at this step of the process; they tend only to report a bivariate comparison between groups on one or two variables of interest and then move on to the remaining steps.

measuring program costs, calculating post release costs and benefits, and computing a measure of the relative costs and benefits to communicate the best policy choice.

In the first step of the cost-benefit analysis, program costs are computed. This usually involves calculating the average "per-day-cost" for a single participant, multiplied by the average number of days served in order to arrive at a final "per person served" estimate of costs (Welsh and Farrington, 2001).[5] In the case of program evaluations in which treatment and control groups are compared, the step ends by computing the *difference* in costs between the two programs. For example, if the treatment costs $5,000 per participant, and the control group costs an average of $4,500, then the realized cost of the treatment is $500 per subject served. This step involves making explicit decisions regarding what is considered a cost when running a program; a decision which can have substantive impacts on findings (Cohen 2005). For example, some politicians view spending on social programs a cost, others see it as a benefit (after all, it employs people; see Boardman et al. 2001: 18-21). The decision of which expenses to label costs and benefits, then, requires an explicit statement of who's perspective you are taking (e.g. a specific department's, a county, a state, a nation, a citizenry, etc.)

The second step involves attaching cost and benefit values to post-program behaviors or events. For example, "costs" may include recidivism events after exiting a program, and "benefits" may involve differences in tax contribution resulting from diverging wages among treatment and control subjects after release (Nagin 2001; Welsh and Farrington, 2001; Cohen 2005). This is by far the more complicated and complex

[5] To be clear, the step also implies a choice of measurement must be made: using average versus marginal costs. Marginal costs are smaller, referring to the added costs of taking on a small increase in services (which usually do not require fixed costs to be altered). In contrast, average costs refer to the full expenditures for each subject served, including fixed operating costs. Either can be correct, depending on the specific question being addressed.

step in the process. It involves explicit decisions regarding which variables are considered relevant (which behaviors to measure), as well as decisions regarding which estimates of costs to attach to those behaviors or events (valuation). In the case of comparing an intervention to a control group, the step involves the generation of an average "per person" cost associated with post-program behaviors for treatment and control groups, representing the sum of cost and benefit information for an average person in each group (i.e. see French et al. 2000; Rajkumar and French 1997; Aos et al. 2001). The difference is then computed, as in step two above. For example, if the treatment group generates an average net cost of $10,000 in post-release behaviors and the control group generates $13,000 in net costs, then the treatment saved $3,000 in post-release behavior costs.

An additional complication in this step involves the handling of selection bias and isolating treatment effects. Most studies incorporate non-randomized designs. One of the most important implications is that comparison groups are nonequivalent, making it extremely difficult to isolate a treatment effect. In these cases, differences in behavior that are due to preexisting differences in risk have to be separated from differences due to program effects through multivariate modeling techniques which require large amounts of data and difficult assumptions to be made. One important method that can be used to combat selection bias is the randomized experiment. The randomization process results in the 'facility assignment dummy' acting as an instrumental variable, meaning that there is no reason to believe that the two groups differ in systematic ways that would also correlate with the outcome of interest (Heckman and Smith 1995). Any post-release

difference in behavior between groups is interpreted as being caused by the intervention in question – not omitted variable bias.

The third step of the general framework involves the computation of a cost-benefit "summary measure." This is the bottom-line information on program costs: how much return should we expect for each dollar invested? All forms of summarizing costs and benefits reflect the difference in costs between the program in question and the comparison group. However, the specific form of the summary measure may change depending on the specific question posed by policymakers and the structure of data available. The two most common are the "cost benefit ratio" and the "net social value" of a policy. I emphasize the second (net social value) because it is a more valuable tool to policymakers than the more common "cost benefit ratio."

The most common approach to summarizing costs and benefits is to create a ratio of the costs and benefit information derived from each program. That is, the difference in costs between the two groups is compared to the difference in post program costs/benefits for the two groups. In the hypothetical example above, this would mean that the additional cost of $500 per participant to run the treatment program generates a reduction of $3,000 in later costs. Thus, the cost-benefit ratio would be 6 : 1, meaning that every $1.00 spent on the program generates $6.00 in savings (relative to the control group).

However, a major drawback to this form of summary statistic is that it is largely ambiguous. For example, consider these two scenarios: In the first scenario, a program cost $1 more to operate per subject than a control but saved $20 in later costs per subject, we would observe an ostensibly large cost-benefit ratio of 20 : 1. However, this would not indicate a substantively large difference in savings when interpreted in scale (i.e.

annual operating capacity). If the program can only hold 700 subjects per year, for example, this would mean the state would save $13,300 per year (i.e. [700 * $1] – [700 * $20]). Now consider a second scenario, in which the treatment program cost $1,000 more to operate than the control per subject, but generated $20,000 in savings per subject. We would still have a 20 : 1 ratio. However, if taken to scale, this would mean that the state would invest $700,000 more to operate the treatment program relative to the control (700 * $1,000), and would save $14 million in later costs (700 * $20,000). This is a net of 13.9 million saved. Knowing that a program has a cost-benefit ratio of 20, then, is not interpretable for policymakers and does not communicate critical information they need. An additional drawback to the cost-benefit ratio is that there are scenarios in which it may take on nonsensical values, such as when two programs cost the same (because this would require division by zero). Likewise, if treatment costs less than a control condition, the mathematical computation of a cost-benefit ratio breaks down and generates non-meaningful values. Although I am unaware of prior literature coming across this situation, it is an event which occurs in this paper (as described in the results section below).

For these reasons, the cost-benefit ratio is far less useful than the alternate "net social value" measure (see Boardman et al. 2001). In this case, the data are summarized in terms of total cost differences, and interpreted both in a per-person metric and at scale (i.e. annual capacity). That is, one computes the total cost of Program A, the total cost of Program B, and compares the difference in terms or a 'per person served' metric, as well as a 'taken to scale' metric. The summary information contained here communicates both

the return on an investment (like the cost benefit ratio), but has the further benefit of allowing that information to be interpreted in a policy context and at scale.

In sum, there has been one general framework for conducting cost-benefit program evaluations which involves four steps. The first is to test whether the cost-benefit evaluation is justified (a typical program evaluation to determine whether there are differences in recidivism or other key variables related to monetizable costs). The second step involves computing costs for running each program (treatment and control). This information is converted to represent the average "per person served" cost generated by operating the program. The difference in costs for the two programs is then computed by subtracting the control group "in-program costs" from the treatment group "in-program costs." In the third step, the post-program costs and benefits are computed for each group (the net value of beneficial and costly "post-program" events) and again translated into a "per person served" metric. The difference in post-program values between treatment and control groups are again computed. Finally, in step four, the information in the prior steps is compared by generating one or more summary measures. The most useful is the "net social value" measure, although the cost-benefit ratio appears more commonly in the literature (i.e. Aos et al. 2001).

Although all cost-benefit program evaluations follow the four steps of the general framework above, there are several different "models" researchers have used to generate estimates of values from the individual steps. For example, what variables are considered relevant "post-program" costs and benefits? Should we only consider costs and benefits directly impacting law enforcement and state budgets? Or, should we include losses to victims? If so, then do you include direct losses only, or pain and suffering as well? Do

we consider the wellbeing of offenders after their treatment programs? Resulting from the complexity of these questions, as well as the divergence of opinions with respect to answering them, several different modeling approaches to estimating costs of post-release behavior have emerged – a topic which is the focus of the section below.

Models in Criminological Cost-Benefit Evaluation

Several models have been adopted by researchers in order to conduct *criminological* cost-benefit evaluations. Each model completes steps 1, 2, and 4 above in the same manner. The key differences in each are found in the post-program estimation procedures (step 3). Namely, in (a) which events are considered costs or benefits, (i.e. deciding whether to include costs endured by law enforcement alone, or to include costs to victims, etc.), and (b) how these events are "valuated" (i.e. the monetary value of an "arrest," the monetary value of a "victimization," etc.). Below, the three most common models used in criminological cost-benefit analysis are described.

1. Criminal Justice System Model. The first (and simplest) approach is to limit analysis to the direct costs endured by the justice system for processing criminal arrests. This is referred to by several names, including the "taxpayer" model by Aos et al. (2001), and the "Cost Savings" model by Greenwood (2001). I use the term "Criminal Justice System" model because it communicates the focus and scope of costs more clearly than prior names.

More specifically, "costs" are limited to losses associated with (a) arresting recidivists (police expenditure), (b) court costs to pursue a prosecution for recidivism,

and (c) reincarceration costs (if returned to prison). By excluding all other potential costs and benefits that may distinguish treatment participants from controls, the model is inherently conservative. The result must be interpreted with caution, as the conclusions are almost certainly wrong; they assume costs are zero for things that we know do actually engender costs (see Cohen 1988). However, its conservative nature does have a benefit. If a program can demonstrate relative cost-effectiveness under the "justice system" model, then statements regarding program benefits are fairly strong.[6]

Of course, this definition of "conservative" has an important caveat. If you believe that state's should and do operate and make decisions based on maximizing their budgets alone, then you are implicitly arguing that victim costs do not matter ; citizens have no standing. This perspective may appear nonsensical when we think about the implications. For example, a state could maximize its budget, since victim losses don't matter, by not enforcing laws. The state would lose nothing (other than victim's losses, which we assume are not important for the state) and gain substantially by not having to spend a dime on police, courts, or corrections. This doesn't sound appealing. Why? Because the goals of the state and law enforcement are explicitly tied to the welfare of citizens. Where does this lead us? This leaves us with the point that victim costs are not zero, and any model which generates a finding that a program is NOT cost-beneficial by leaving these costs out must be wrong. More specifically, it leaves us with two statements. (1) If a program demonstrates that it is cost-beneficial in terms of a criminal

[6] There is a potential exception to the assertion that the conservative nature means more defensible conclusions. It would not be true if there are *perverse* non-crime-effects of a program which overwhelms criminal justice savings. For example, if a program generates lower recidivism – but also generates high rates of deleterious behaviors not measure as a cost to the state (i.e. unemployment, physical health deterioration, or death). It is unlikely that such a case would occur because these analogous behaviors tend to correlate positively with offending, not negatively (Gottfredson and Hirschi 1990).

justice model alone, then we have strong evidence that the program is a good one – because it has won out despite an unrealistically unleveled playing field. In contrast, (2) if a program does not demonstrate cost effectiveness in terms of a criminal justice model, we may conclude very little. In short, the results would absolutely be wrong in terms of citizens' losses in both cases. But the direction of the bias is toward null findings so we may use this information in one case only (a rejection of null findings) but not the other.

For example, Aos et al. (2001) used this model within a meta-analysis framework to examine the costs and benefits of ABE (adult basic education) among incarcerated adults in the state of Washington. They found the average cost for the programs were $1,972 per subject served.[7] Drawing on three published studies, they computed an effect size of -0.11 with regard to recidivism. Drawing on their specific criminal justice "average costs" of recidivism for Washington, they estimated the state would observe $1,852 fewer costs associated with later offending relative to a policy option of not providing this service. Therefore, the authors concluded the ABE programs were not cost effective. Again, this was based on assumptions which create very low estimates of the value of crime. No consideration was given to pain and suffering of victims, financial or medical losses of victims, or other behavior (i.e. welfare use, employment changes, mortality, etc.).

Another example (which was slightly more complicated) can also be found in Aos et al. (2001); an example in which they reviewed the cost and benefits of adult boot camps. They again drew on published evaluations and computed effect sizes for recidivism impacts across the studies. Identifying 11 evaluations meeting their minimal

[7] All figures were translated into 2005 dollar values in order to facilitate comparisons. For more details on conversions, please see Federal Reserve web-site at: http://woodrow.mpls.frb.fed.us/research/data/us/calc/

qualifications regarding research design, they found the average effect size was zero. That is, adult boot camps had no average impact on recidivism. However, they found that adult boot camps could still be cost-effective, depending on how they were used – because they were cheaper than alternative prison terms (due to the shorter duration of sentences assumed).

Drawing on MacKenzie and Piquero (1994), they argued that boot camps vary in their cost-utility because they vary in how they draw their population (i.e. divert inmates from traditional prisons, or net-widen to offenders who would otherwise have been on probation). Therefore, they produced several estimates of the costs and benefits of adult boot camps under varying conditions of implementation, using hypothetical estimates of the proportion of the boot camp inmates that would be "true diversions from prison" (meaning that money would be saved even with identical recidivism rates) versus "net-widening" approaches (meaning money would be lost). If an adult boot camp contained 100% prison diversions, they estimated that savings would be $9,725 per offender served. If the camps only served 75% prison diversions, then they estimated camps would save $3,500 per participant served. The monetary estimates were wholly derived from the difference of in-program costs, since there were no differences in post-program recidivism in their models.

2. Direct-Cost Model. A second model used to complete the general framework is the "direct-cost" approach (see Aos et al. 2001; Schweinhart et al. 1993). This model utilizes the same recidivism estimates as the "justice system" model, but adds in limited information on victimization costs. Here, the limit is to "direct" measurable costs, such

as the amount of money stolen from a victim, the value of hospital fees for an assault or rape victim, etc. In this case, the scope of the "post-program" costs and benefits are expanded to variables which are clearly and defensibly measured in monetary terms. Thus, the distinguishing characteristic of this model is that it conceptualizes the relevance of an intervention beyond the state budget per se, and includes the direct costs to citizens as equally relevant.

For example, Aos et al. (2001) replicated their earlier "justice system" models using the "direct-cost" model for both programs listed above. By adding in the direct losses to victims (averages) in the above studies, more monetary "weight" was given to recidivism. In the case of prison-based education (ABE), this led to a reversal of their earlier findings. Instead of the state losing $120 for every subject they serve through the program (i.e. $1,972 costs - $1,852 crime reduction), they now estimate a net benefit. That is, they estimate $9,176 of direct losses avoided for every $1,972 invested, a benefit of $7,204 for every subject served.

In contrast, the inclusion of the "victim's perspective" had no impact on their estimates for Adult Boot Camp from above. Again, because there was "zero" difference in post program behavior for the boot camp participants, the contribution of post-program weighting procedures has no impact. In this case, no matter how recidivism is monetized, the difference is still 'zero.'

Turning to another example, Austin (1986) evaluated the costs and benefits of an incapacitation strategy for the state of Michigan. In this analysis, Austin calculated the difference in costs (per-inmate incarceration costs) for the "old system" (traditional sentencing patterns) versus the "new system" (incapacitation driven sentencing changes).

Here, program costs were different because the average time served would change, and more 'light' offenders would be sent to prison rather than intermediate sanctions. Using estimates of recidivism under both systems, he concluded recidivism would be reduced if the incapacitation approach was adopted. However, by incorporating "criminal justice costs" for recidivism, as well as "direct victim losses," he found the incapacitation strategy was not cost effective. That is, the expense of housing inmates for the additional time would cost far more than the savings generated in terms of criminal justice expenditures and direct loses to victims.

3. Full Cost Benefit Model. A third model for estimating the costs and benefits of criminological interventions is the "full" cost-benefit model (Greenwood et al. 2001), and is considered the 'gold standard' of the field (Nagin 2001). In this case, the information in each of the two prior models is included in estimates of post-program effects. However, additional information is included regarding (a) additional outcome measures (such as social service use, life expectancy changes, wage changes) if possible, as well as (b) intangible costs of crime. The latter addition is the most abstract component of crime-value estimation. It refers to the monetary values that econometric estimation procedures have placed on various criminal acts. For example, the pain and suffering of a rape victim, or the value attached to "fear of crime" in a community when a burglary occurs. These are all real, substantive, and important costs associated with recidivism (they are events that people would pay to avoid – indeed, people do pay real money to avoid). If a researcher's goal is to estimate the value of an intervention to citizens, then it makes sense to include these costs. To be clear, some are reluctant to do so because the abstract

nature of the data, the complicated mathematical models on which estimates are based (making some unsure of the reliability of estimates), and the difficulty of supporting assumptions required to generalize estimates to a particular data set in question (see Greenwood et al., 2001). However, the quality and reliability of estimation procedures has increased remarkably in the past two decades, and is widely accepted and used in nearly every analogous field to criminology (i.e. environmental science, public health, military science, transportation policy, etc.; see Cohen 2001, 2005). As a result, modern cost-benefit researchers should be reluctant to ignore intangible costs when producing program evaluations.

Nagin (2001) is particularly adamant in his support of this latter model, arguing that the broader social costs of interventions are critically important to good policy making and accurate assessment. For example, he applauds Hawkins et al. (1999) for demonstrating that a school-based prevention program not only reduces offending, but also reduces sexual promiscuity and improves school performance. Likewise, he commends several classic evaluations (i.e. Perry Pre-school Project, and Nurse Home Visitation) which have shown treatment to not only reduce offending, but also increase home ownership, increase high school graduation rates, as well as employment and wages, etc. (see Olds et al. 1988; Schweinhart, et al. 1993). Collecting outcomes other than recidivism does require criminologists to venture outside of their comfort zone, and to tackle methodological issues regarding data collection and use from analogous fields. Despite the effort inherent, it is important in that many of these latter measures may lend themselves well to monetary interpretations and thus facilitate not only a "more comprehensive picture" of the impact of evaluations, but also increase the accuracy of

cost-benefit research. In contrast, many criminology studies have followed the Aos et al. (2001) method of limiting analysis to criminal justice outcomes only. This latter method is endorsed by some researchers because of the familiarity of the data, the conservative test implied, as well as the parsimony in the comparisons. But it is an approach that worries Nagin – because the limited scope of the analyses may preclude important information from coming to light. In sum, the "full cost benefit" model includes criminal justice losses, direct victim losses, and intangible victim losses resulting from recidivism. It may include other non-crime costs as well, if possible (i.e. such as changes in wages, welfare use, physical health, etc.), and is considered the most useful and important of the models reviewed.[8]

Examples of this model are prominent in the literature. Cohen 1988 reanalyzed the above study by Austin (1986) but added in the "intangible" costs to victims and communities – a change which reversed the conclusions of Austin (1986) described above. For example, in the prior analysis by Austin, the monetary loss endured by a victim of "rape" was approximately $410 in 2005 dollars (i.e. the value of hospital visitation and direct injury / property damage during the average rape).[9] The value does not ring true – the very point Cohen is trying to make; there are losses to victims that are not captured in direct-measure methods. Adding in econometric estimates of

[8] The literature's emphasis on recidivism as 'the' key measure, while other social behaviors are considered important but secondary is made for two reasons. First, policymakers deciding between criminal justice programs have a substantive interest in grounding their decisions on deployment of funds in terms of crime-related outcomes; it is what they and the public expect. In addition, it is a practical consideration. The empirical literature on program evaluation has demonstrated that recidivism routinely accounts for 80% or more of the total monetized-utility of a program, and other measures often only have a secondary impact on findings (see Barnett et al. 1985). Testing this assertion, Logan et al. 2004 employed a large number of potential outcomes of a criminal justice intervention, including traffic accident changes, mental health visits, hospital visits, welfare use, as well as crime. In their data, the crime measures drove their analyses, and the other measures contributed little in terms of their ability to choose between programs.

[9] Conversions of monetary values to specific dollar-years conducted according to procedures outlined at http://woodrow.mpls.frb.fed.us/research/data/us/calc/

pain/suffering, and reductions in wellbeing of victims, Cohen (1988) estimated the value of a rape to be $92,673 (2005 dollars). In a similar fashion, he drew on econometric literatures to adjust other crime types in the data set, generating more complete (and thus higher) estimates of crime costs. Using these estimates, he found that Austin's prior conclusions were reversed. That is, the value of the crime reduction associated with incapacitation sentencing policy would indeed save citizens of Michigan more money than the "old system" of sentencing sans the incapacitation philosophy. The difference in the two studies was entirely due to how they "weighted" offending.

Summary: Cost Benefit Models

In summarizing the difference between model choice, it is helpful to use taxonomy to clarify the differences between competing approaches. This taxonomy divides "costs" of crime into four dimensions, by classifying "costs" as either (a) primary or secondary, and (b) tangible or intangible. The term "primary" costs refer to losses endured by the victim involved in the criminal event. Term "secondary" costs refer to the costs endured by other specific individuals outside the event (such as family members who suffer pain when a child is killed), or social broader social costs (such as neighborhood disinvestment that results from crime events).

The specific costs within these two categories can be grouped as either "tangible" losses or "intangible." Tangible losses refer to easily measured events in which monetization is clear and 'direct;' there would be no substantive debate surrounding the valuation of the event in question (i.e. the amount of money stolen in a robbery). The term "intangibles" is a misnomer, in that they are "difficult and abstract" events as far as

monetization and measurement is concerned; not truly "immeasurable" as the term implies. This would include "fear of crime," "emotional duress," and "pain and suffering" among primary or secondary groups. Consider Figure 1 below to clarify how losses are conceptualized within this taxonomy for a hypothetical "robbery." The specific losses itemized are examples of costs, but not an exhaustive list with reference to robbery. In short, the models discussed above differ in which cells they allow to 'count'. For example, the "Justice System" model only allows cell B to count. The "Direct Losses" model only allows cells A and B to count. The "Full Cost-Benefit" model must include cells A, B, C, and D, and attempts to contain non-criminal justice expenses as if possible.

[Insert Figure 1]

In sum, there are three models used to approach the "general framework" of cost-benefit evaluation within criminology. They are all similar in that they follow steps 1 and 3 of the general framework in the same manner. However, they differ in terms of post-program estimation procedures regarding costs and benefits (step 2). That is, (a) what is considered a relevant behavior or event, and (b) how to compute the value of those behaviors or events. The choice between models is derived from a comprehensive consideration of emerging research methodology within the criminological cost-benefit literature, and the needs or interests of policymakers. Given that there is still some debate about the adequacy or reliability of these competing choices, the most reasonable approach appears to be the replication of an analysis using each of the three models so

that full information can be conveyed and the impact of model choice can be explicitly communicated – an approach adopted in this dissertation.

Having reviewed the general cost-benefit framework and the extant models being used to evaluate programs, I now turn to a discussion of the process by which costs are attached to events; a process referred to as "valuation." The goal of the valuation literature is to generate robust estimates of the cost of events, costs which may be used as "plug in" values for researchers conducting program evaluations (Boardman et al. 2001). The following section reviews the monetary values used in the literature for various recidivism component-cost (i.e. cost to the justice system, the victims, and society). Facilitating this presentation, the valuation literature is separated into "criminal justice" components, followed by "victim" components. The section then concludes with a discussion of different approaches which have been used to put the "valuated" components together. That is, the different assumptions people have made with respect to event occurrence or the level of detail needed in order to attach values to those events; differences in what is assumed constant versus measured across subjects. In both cases (computing components and combining them) there has been substantial variation in the published literature.

Valuation of Recidivism

1. Computing Components: Valuating Criminal Justice Expenses

Calculating the costs of crime to criminal justice agencies involves at least three major sources of monetary loss. These include (1) costs of the investigation and arrest event, (2) costs to the courts for processing the event, and (3) costs associated with

administering sanctions. Each is discussed in turn, with emphasis placed on the difference between "top-down" and "bottom-up" methods to valuation. All values reported below have been converted to 2005 dollars, unless otherwise noted.

Recidivism generates a variety of potential costs. In reported crimes, costs will always involve costs to police who respond to crime scenes, investigate offenses, make arrests and file paperwork. In some cases, this may lead to emergency response to crime scenes, booking at local jails, and intermediate detention. Various other agencies may endure costs as well, depending on how a case proceeds. If a case is prosecuted, then losses are endured by the courts (prosecution, judges, and administration). Detention may continue through the course of prosecution, or not. Court events themselves may be long lasting and involve many court visits, or may end after a single hearing. If a recidivism event leads to sanctions, costs may vary depending on whether or not sanctions involve jail time, a prison return, or intermediate sanctions. In short, there are a number of avenues through which recidivism events may generate losses to the state. This implies that researchers should articulate which events occurred amongst their subjects and attach monetary values to these events, an issue I turn to in the section below.

Making an arrest. There have been two general methodologies used to valuate criminal justice costs, including costs of arrest to police. The first is a "top down" approach in which aggregate budget data (for a county, a state, or a nation) were regressed on crime rates to estimate the proportion of police budgets attributable to arrests, or each type of arrest. The second approach has been a "bottom-up" method in

which actual arrest events were examined and resources deployed are recorded (i.e. number and hours of police officers, ambulance use, etc.). Again, this may be across arrests 'in general,' or computed for different crime types.

The "top-down" method is more commonly employed when valuating criminal justice costs, although there is important variation across studies with respect to the level of detail in the data used. The simplest methods have involved taking a national estimate for police expenditures and dividing the total by the proportion of all arrests that fall into different crime categories. Essentially, this means a budget total is divided across the proportion of arrests made for different categories of offenses (see Harwood et al. 1984, and Rice et al. 1989).

A typical example is an evaluation conducted by the Metropolitan Dade County Department of Justice Assistance (see Cohen et al. 1994: 127-8). Reviewing data from 40 jurisdictions, they found a total of $7,713,658 was spent on booking annually, spread out across 41,572 cases. This resulted in an average of $186 per booking event ($320 in 2005 dollars). Using similar data, they estimated an average arrest in their county (including investigation) ran $820, bringing the total cost of arrest, investigation, and booking to $1,140. However, this figure is limited in that they generated a single dollar value for all crimes, rather than computing estimates of costs by crime types. Likewise, it only speaks to one state, which may limit the generalizability of findings.[10]

Correcting for these deficiencies, Rajkumar and French (1997) used DOJ data on national-level expenditures for police (annually), combined with estimates of the national count of arrests made for each of 10 offenses (annually) to compute police expenditures

[10] An additional limitation, of course, is that the data speak to a specific historical period (the mid 1980s). As technology and offending patterns change, the costs are expected to change as well (i.e. the introduction of automated booking systems, DNA technology, etc.).

by crime type.[11] Unfortunately, the authors did not report the actual estimate for the cost of arresting offenders. Rather, they only reported the summed total of the average police cost, the average court cost, and the average prison cost all rolled into one figure for each of the crime types. Although they did report the final monetary estimate for each crime type for "criminal justice costs" in whole, their publications do not allow readers to capture their estimates of the cost of making an arrest.[12]

In a more recent study, Aos et al. (2001) used a similar method as Rajkumar and French, although their data were limited to a single state (Washington), and they used fewer crime categories. Specifically, they regressed the state's police expenditures on the rate of arrest across four crime types (violent felonies, violent misdemeanors, misdemeanors, and traffic arrests). They concluded that the average "serious violent arrest" (i.e. homicide, sex offense, robbery, aggravated assault) consumed $16,084 in police resources. In contrast, the average property crime consumed only $2,421 in police resources.[13] Although these estimates broke "crime" into more categories than some top-down valuations, the categories still leave something to be desired. For example, they were built off the assumption that a homicide, rape, and aggravated assault case all use the same amount of police resources.

[11] These offenses are: robbery, burglary, drug, prostitution, assault, motor vehicle theft, theft, forgery, fencing, and gambling. Interestingly, they add homicides into the offense category describing the motive for the homicide. Thus, a homicide resulting from a robbery is classified as a robbery, creating a weighted average of resources deployed for homicide-robberies and non-homicide robberies.

[12] Unfortunately, the original document cited as containing these component figures and methodology details is unpublished, and the original holding agency (RTI) reports that the document has since been lost (no electronic or paper copies could be found). The authors have not yet responded to requests for information or copies of their paper.

[13] The models used by Aos et al. (2001) do not provide detail as to the treatment of crimes which do not fit into categories they coded for. For example, what happens with drug-felonies? How about property-felonies?

In summary, the top-down method is the more common approach to valuating police expenditures for arrest. Usually, this includes a model in which the police budget on the left-hand side and number of crimes is on the right-hand side. Thus, the approach generates an estimate of the marginal cost of 'one more' more. However, the ease of computing these estimates has come at some cost. First, the models may have specification problems, in that there may have been omitted variables that both correlated with the number of crimes observed and the police budget. For example, cities with larger populations may have higher police budgets due to higher tax bases to support them, and more crime as a function of the urban nature of larger cities.

A related, and perhaps more important problem, is one of endogeneity. That is, it is unclear whether increases in arrests lead to larger police budgets (the premise of these models) or whether increased police budgets lead to larger numbers of arrests. For example, the crime rate may stay constant over time, but increasing the number of officers or technology of officers (i.e. budget increases) may lead to more arrests within that 'constant' pool of offenses. The literature suggests both events occur, and the latter scenario is far from negligible (see Skogan et al 2004).

In contrast to the methods above, a handful of authors have attempted to examine actual arrest events and build estimates of crime scene costs from the "bottom-up." Although far more labor intensive than "top-down" approaches, the estimates derived from real events are likely more accurate because they rely on fewer assumptions, and reflect the marginal cost of an event. Estimates derived from bottom-up methods generally imply a lower cost associated with arrests.

For example, in their evaluation of the Anne Arundel Drug Court in Maryland, Crumpton et al. (2004) relied on phone interviews in which police chiefs estimated (1) the per-hour cost for an average street officer, and (2) the number of hours of police time involved with a typical arrest (i.e. travel to crime scene, make arrest, travel to booking station, book, fill out paperwork, and transfer custody to detention center). They reported that the average cost of arrest to police was $508.42 (see p. 30 – 32). This estimate referred to all arrests, regardless of crime type. The method generated a smaller estimate of the "average" arrest found in any of the top-down methodologies. However, it is still limited in that arrest costs were assumed constant across types of offenses, an assumption that is difficult to justify.

A similar approach was used in the Perry Pre-school Evaluation (Barnett 1985; Schweinhart et al. 1993), in which police costs were estimated for arrest processing. Interviews with officers in Ypsilanti revealed the typical arrest took between 1 – 4 hours. Thus, the authors created two figures representing either a single hour of police salary (the petty arrests) or four hours of police time (a serious arrest). All arrests were classified as serious or petty, and these values were attached accordingly ($56 for a petty arrest, and $213 for a serious arrest). Again, these estimates were derived from salary estimates only; they did not include fringe and assume one officer per arrest.[14]

Improving on this framework, Cohen, Miller, and Rossman (1994) built up estimates of police expenditures across crime types. Here, they used surveys tapping the

[14] Importantly, the Perry evaluations excluded (a) traffic offenses and (b) homicides. Both events occurred in their sample, but the authors felt that the traffic offenses were too petty to be considered crime, and the offense of homicide was too substantively unique an offense to be included as a crime their program should be expected to impact. Likewise, they were concerned that the expensive nature of homicide would overwhelm their estimates if included (see Barnet 1985: 57).

number of hours required to process arrests and crime scenes (across crime types), multiplied by an average hourly wage of police. For example, they found the average robbery scene took 1.5 hours to process, the average murder took 3.5 hours, the average rape required 2.4 hours, and the average assault used 0.9 hours of police services. Finding that the average police resources were worth $45.05 per hour, murders cost the police $198, rape costs $34, robbery costs $86, and assault costs $52. Importantly, these estimates referring to processing a crime scene and responding to a crime scene, they do not include the cost of investigation (a cost which may be negligible for some offenses, but exceptionally high for others).

Building on Cohen et al. (1994), Miller and colleagues (1996) expanded their analysis to include more offenses and updated data sources. The authors used survey estimates of hours required to process a scene across crime types, and refer both to police services to investigate and arrest, as well as alternate resources such as fire/rescue deployment. Here, the authors found the average Homicide ran $1,810 in police costs, Robbery consumed $487 to process a scene, the average Assault consumed $238, and the average Motor Vehicle Theft (MVT) required approximately $223.[15] Importantly, these

[15] The police-costs in Miller et al. 1996 are reported in per-crime metric, including crimes not reported to police (which would have no costs if our point of reference was to valuate crimes in which an arrest occurred). Therefore, data on the rate of reporting victimization derived from the NCVS was used to convert their estimates into a "per reported crime" metric. The estimates were not further converted, however, into a "per crime cleared" metric, because this conversion would require an assumption that the money identified by Miller and colleagues was exhaustively spent on events in which an arrest occurred, rather than being distributed across events in which an offender is caught or not. For example, their estimate of the cost of a rape is $37 in police expense per event. Using the most proximate NCVS estimates (1996) the data show that 55.7% of victims report their rape to police, which means that the *police* expense is actually ($37 / 0.557) = $66 per investigation. However, although the show that only 55.2% of reported cases are cleared – I can not convert this to the per-arrest metric of ($66 / .552) = $120 per arrest, because that conversion would assume that police spent no money on cases with no arrest, and all of it on cleared cases; an unlikely assumption. A better approach would be a conversion based on the difference in amounts of resources spent on the average cleared and uncleared arrests (i.e. a ratio), but none exists. In this case, it is plausible to assume that expenses are roughly similar across cleared and uncleared cases, which means that the average of $66 is the most defensible measure.

estimates were slightly higher for violent crimes than other bottom-up figures in part because they include the full crime scene costs (i.e. fire rescue services that may respond). However, they still did not include the cost of investigation to police.

These estimates are strikingly close to estimates derived almost a decade earlier using a similar methodology. Larson 1983 examined activities of police during arrests across a variety of crime types (i.e. taking physical custody of an individual and completing required paperwork). He estimated a robbery consumed around $176, an assault consumed around $148, and a MVT cost around $125 (see Austin 1986).

Finally, in the most recently updated estimates of the bottom-up costs of crime, Cohen 1998 re-estimated the cost of investigation for seven crime types. This final paper included all the above information from the earlier papers, but added in the costs of investigation. Thus, we see a large jump in the cost of costs of crimes with particularly expensive investigation efforts associated (i.e. rape and homicide). Here, he estimated the cost of a homicide was running $10,614, the average rape investigation cost $3,800, the average robbery cost $869, and the average larceny cost $264 in police expense.

In short, the bottom-up estimates are far more accurate than top-down methods, as they rely on observation rather than assumptions regarding the allocation of police resources. In addition, the bottom-up studies of Miller et al. (1996), as well as Cohen et al (1998) included costs endured by rescue workers in addition to police responding to a crime scene. The magnitude of the error involved in taking the "easier" approach to valuation (the top down approach) is rather impressive, as the difference between the costs of a robbery is $16,000 in the Aos et al. (2001) report, whereas the actual cost is likely closer to $869 (as found in the Cohen et al. 1998 study); a difference of around

1,800%. Again, the reason for the difference is that all money that was spent on infrastructure, police salary and administration, and capital or operating costs associated with prevention or mundane policing (i.e. not making an arrest) was assumed to be "zero" in the top down methods. Instead, all those resources (which represent almost the entirety of police expenses) were attributed to arrest events.[16] See Table 1 below for a summary of "cost of making an arrest" estimates across the field.

[Insert Table 1]

Court Processing. Just as with the above literature, two approaches have been used to generate estimates of the cost of courts among arrestees. The top-down approaches vary in their complexity, but all rest on assumptions regarding the allocation of budgets exhaustively across variables included in models. There are a growing number of bottom-up valuation studies emerging in the literature. At least in part, the growth appears to be due to the increasing number of drug court evaluations in which court costs must be estimated within the bottom-up framework when computing differences in treatment costs.

[16] As will become more clear in the "victim losses" sections below, the valuation literature has been far better at using diverse and creative tools to estimate the cost of victim losses, rather than the costs to criminal justice actors. This may be due, in part, to the larger impacts that victim losses are likely to have, and that audiences (i.e. policymakers, advocacy groups, etc.) find the research on victim losses more interesting or intriguing than similar studies on 'policing losses.' This is problematic, because it leaves us with less satisfying impressions of reliability in the criminal justice component valuations. A plausible direction for the future of cost-benefit criminology, then, may be to undertake creative approaches to policing valuation, as with victims. For example, researchers could compare the rate of injury among officers who are engaged in making arrests (across crime types) versus normal patrol. The marginal increase in physical harm (or property loss, such as a wreck with a police vehicle) while engaging in an arrest could contribute to our estimation of police losses associated with making an arrest – losses which likely vary by crime type and are worth valuating.

The Dade County Department of Justice Assistance again used a top-down method to estimate the costs of prosecution to courts (see Cohen et al. 1994). This included an estimate of $15,694,845 spent on arraignments of 42,629 cases, for a total of $368 per arraignment (or, $633 in 2005 dollars). Likewise, the average Pre-Trial hearing was estimated to run $2,137, a Jury Trial ran $51,889, and a Sentencing Hearing ran $607. The added cost of a jury trial may seem dramatic, but recall there are jury fees to be paid, added security, and substantially more court room time and case preparation. In addition, jury trials are often only used in the most complicated or heinous cases, implying they may take more resources to process. These figures can be used to compute costs recidivists generated; regardless of what crime type they were charged with. In contrast, most studies estimate the cost to courts by crime type – by drawing on average costs associated with processing different types of charges.

For example, Aos et al. (2001) used a multivariate regression model to estimate the value of court costs in the state of Washington. Their model examined variation across counties, but limited the budget lines to those expenses directly associated with criminal prosecution. This total was regressed on the count of crime type *convictions*, including (1) homicide, (2) all other violent felonies, (3) all other non-violent felonies, and (4) the number of "non-criminal" court filings (the only control variable). They used the resulting equation to predict the cost of the average court expenditure by crime type (a point estimate for each crime). They found, for example, that the average homicide conviction cost the local courts $97,034, the average robbery cost $18,399, and the average property crime cost $1,675. The model had several strengths. First, it did not assume costs are equal across all criminal prosecutions. Second, they excluded "non-

crime" cases, rather than assuming these funds were spent on criminal cases. These represent stark improvements over simpler top-down methods. However, the model was still limited. For example, it assumed that costs were only spent on convictions; that the cost of a non-conviction case was zero. That is, they assumed the entire budget examined was only spent on cases which return legal convictions, and none on cases which resulted in acquittal, diversion, informal possessing, or other administrative and interagency work. Inasmuch as any of the budgets from the courts actually funded work other than action directly involved in the conviction of a case, then these estimates would have been biased upwards.

Although they used a bottom-up method to estimate arrest costs, the Perry evaluators turned to a top-down approach to estimate court costs. Here, they estimated the costs for juvenile and adult court separately. Within each of these domains, they used the "number of cases referred" as the divisor for that domain's total court budget (i.e. computed separately for juvenile and adult cases). This gave them the proportion of the budget each case presumably used. For example, there were 808 referrals to juvenile court in one of their years, generating a cost of $647 per case in 1981 dollars (see Barnet 1985: 46). There was less bias than in other top-down studies because costs were divided across all cases which appear (referrals), rather than dividing by number of convictions only. The assumption, of course, was still fragile; implying cases which resulted in dismissal cost the same as cases which go to trial. Regardless, there was less bias here because the diversion and dismissed cases were allowed to soak up some court costs, rather than assuming they cost nothing to process.

In contrast, several studies have attempted to use a bottom-up strategy to compute court costs. For example, Crumpton et al. (2004) used county level budget data to construct the average hourly cost of judges and district attorneys (as well as an hourly-metric of overhead costs) in Ann Arundel County, Maryland. They used personal interviews with agency officials to derive the average number of hours used in a prosecution by each agency. Using these estimates, they computed the "average" cost of a traditional prosecution to attach an estimate of the cost of prosecuting the average arrest. This would represent, of course, the average number of hours across all crime types and all court events (i.e. plea versus jury trial). Unfortunately, the authors did not report any of the component parts of their calculation. Rather, they simply reported following these steps to attach costs to recidivism in their program evaluation. They did not tell readers the cost per hearing, the number of hearings in a case, or any descriptives on the number of prosecutions or outcomes. The problem, then, was that their analysis and findings were opaque; not allowing a reader to capture basic information by which to replicate or evaluate their conclusions. It also means that basic cost elements can not be used as "plug in" values for other researchers.

In a comprehensive study of criminal justice costs, Cohen et al. (1994) used national data to generate two estimates of court expenses for a number of crime types (based on conservative versus relaxed assumptions). First, they used data on the average number of hours taken to prosecute a variety of court cases (i.e. murder, rape, robbery, and assault). This included time in court, conferences, and research. They combined this with national data on the average salary of court workers involved in prosecution to compute the national average cost of prosecuting a variety of offenses. For example, they

found that the average homicide required 13.2 hours of time from each worker involved, the average robbery 5.1 hours, and the average assault case 3.7 hours. Using these figures, they estimated the average homicide cost $945, the average robbery ran $516, the average assault cost $430, and the average rape cost $602. However, these estimates were a "lower-bound" because they did not account for the overhead and other expenses caught up in processing court cases. Thus, in their second estimate, they used national estimates of the costs to the courts which included overhead as well as salary data, which came to $6.88 per minute. Using these "upper-bound" estimates, the costs increased, with a homicide running $6,361, a robbery $3,567, an assault $1,805, and a rape running $2,075. The difference may seem dramatic, but recall that the later method accounted for overhead as well as the secondary employees involved in case processing (i.e. clerks, bailiffs, and support staff). A limitation of both estimates, however, was that they represented averages *across all court outcomes* (dismissed cases, plea bargains, and trials). As such, they are of limited use as "plug in" values for program evaluations which attempt to valuate recidivism with known court outcomes (i.e. plea versus acquittal).

Taking an approach more amenable to program evaluation, Czelen, Wheeler, and Mott (1978) constructed a detailed assessment of the use of court time across a number of types of court events for Brooklyn, NY. Based on hourly wage data, as well as a recording of the number of hours and personnel present in different types of hearings, the authors estimated the cost of four types of court events. They found the average arraignment ran $655, the average felony hearing ran $1,646, the average minor hearing (i.e. motions and misdemeanor hearings) cost $562, and the average sentencing hearing

cost $911. Although this study is 25 years old, a recent assessment of the valuations by Zarkin et al. (2005) found the estimates were still accurately capturing court resource use across prosecutions in Brooklyn, NY. Note the similarity of these estimates of hearing costs with other estimates listed above (i.e. Dade County Department of Justice Assistance). Again, a key relevance of this methodology is that it allows a cost to be attached to hearing events, without having to make assumptions about the court outcome (plea versus trial, etc.) or about differences in court use by crime type.

An important caveat is that the above studies did not account for public defender use. Indigent Defense consumes a large amount of tax dollars ($991 million nationally in 1986, which is approximately $395 per case in 2005 dollars; see Cohen et al. 1994). For example, the Public Defenders Office in Maryland handled just under 175,000 cases in 2004, for a total cost $61.8 million (see Maryland Legislative Budget, 2005). This reflects an average of around $353 per case, or $365 in 2005 dollars. This is not far from the national average of $395 cited in Cohen et al. (1994), after converting from 1987 values to 2005-dollars.

However, there has been little research attempting to break down indigent defense into component costs (i.e. cost per charge type, or cost per hearing type). Using the nationwide data, Cohen et al. 1994 estimated that indigent defense cost approximately $98 per hour, derived from salary and fringe estimates. Using information on the average hours to research and handle a case, this implied that cases *which have public defenders* cost $1,289 for a homicide defense, $499 for a robbery defense, $364 for an assault defense, and $713 for an average rape defense. These numbers represented lower-bound

estimates for indigent defense; they did not include overhead and were averaged across all court outcome types.

A final caveat is that many arrested individuals spend at least some time in jail during their court process. This is particularly true of parolees, the subjects of the present study. This could include a single overnight stay, or a stay which endured through the entire court process. To capture this, studies generally use the "day-cost" of detainment computed from either a national or statewide estimate. (The day-cost is usually derived from an annual budget divided by the average population count for that facility.) The major difference between program evaluations tend to be in whether they actually measured the use of jail among their subjects, or impute the "average" jail use from national or state estimates onto their subjects (i.e. detention averages based on offense type).

For example, Crumpton et al. (2004) observed the days detained for each recidivist in their sample, which they multiply by the "per day" cost of jail in that county. This generated the cost of "detention during trial." In contrast, Aos et al. (2001) used several estimates combined together to generate a cost of detention during trial. They use (a) the average day-cost of jail, which is multiplied by (b) the estimated average length of detention for each crime type during trial. For example, they estimated the average Robbery case generated 0.29 years in detention during trial. In contrast, they estimated that the average drug charge resulted in only 0.19 years of detention while processing through the courts.

Sanction Administration. In a portion of the recidivism events, a sanction will be administered (i.e. prison time). Most research accounts for this by computing an average daily cost for prison across a specific state. Others differentiate incarceration costs by computing a figure for jail costs as well (which are often slightly higher than prison costs). If incarceration is less than one year, the jail cost is used. If sentences are greater than a year, the prison rates are used. However, there are important differences in decisions researchers have made with regards to the length of sentences which will be served. And how much incarceration is assumed to cost.

For example, in Crumpton et al. (2004), the authors computed day costs for both jail and prison, and attached those estimates to observed sentence lengths resulting from sentences for new crimes, or new terms of parole. This research was not clear as to whether maximum sentences were used, or sentences after removing 'suspended' time – a difference which can be substantial in terms of costs.

Aos et al. (2001) were more explicit with regards to this point. They computed the average cost of a day in prison for the state of Washington, which they attached to average sentences handed out for their crime type categories (as described above). Further, they used published reports on sentence duration across crime types for the state of Washington to adjust for early release; generating an estimate of actual time subjects would likely serve. For example, they estimated that the average robbery term in their state would be 5.7 years, of which offenders would serve an average of 4.4 years (77.19%). The average "property" offense resulted in 2.4 years of prison time, of which only 1.7 would be served on average (70.83%). Thus, they attached their estimates of prison day costs to these adjusted "average time served" values when valuating sanction

costs. A major limitation of this research, which was particularly pronounced in terms of sentence lengths, was the grouping of charges together which were likely very different from one another (i.e. homicide and aggravated assault were assumed to have equal average terms in their design).

Taking a different approach to sentence lengths, McCollister et al. 2004 drew upon a randomized experiment focused on prison based therapeutic communities. They examined the "number of days incarcerated" at five years after release. They found the treatment group average 544 days of incarceration, whereas the control group averaged 626 days (a difference of 81 days, which came to $5,265 dollars worth of 'difference,' assuming prison costs around $65 per day). However, these findings should be taken with some reservation, because the methodology employed may have suffered from censoring bias in the sanction measurement.

For example, what if the actual sentences handed down were the same between groups, but the treatment group took longer to be arrested and processed (i.e. offended less often or delayed offending for a slightly longer time after being at risk)? In this case, the control group would be well into their prison terms at the 5-year cut date, and the treatment group would just be starting. The censoring of data at 5-years would mask this; creating an illusion of program savings. In fact, the treatment group could have been sentenced for more time on average than controls and we'd miss this as well because of the censoring problem inherent in their measurement choice (again, assuming late start among the treatment group). For example, suppose Panel A of Figure 2 represents the true total sentence lengths the treatment and control groups were ordered to serve (the two horizontal lines are equal in length). In this case, Panel B of Figure 2 represents the

artificial censoring invoked by the 5-year cut off date; the lines appear to be different in length. The point here is that the potential error is easy to avoid; researchers need only measure sentences as ordered by the court, rather than "days served" by a cut off point.

[Insert Figure 2]

Interestingly, all of these studies fail to appreciate that different sentence types likely result in different locations in which offenders serve their time. The average drug possession inmate sentenced to three years in prison will serve time at a minimum security prison. The average murder inmate doing 30 years will serve their time at a maximum security prison. The per-day costs here can be significantly different.

In sum, valuation methodologies across criminal justice expenditures tend to follow one of two methods: top-down or bottom-up. The top-down are more common, but often rely on untenable assumptions. Likewise, the actual valuation estimates produced are often ill-suited for program evaluations, because they lack specificity in costs across crime types, etc. The bottom-up methods make fewer assumptions about budget allocation and generally produce more defensible estimates of the actual costs associated with handling a case at each stage of the justice system. The bottom-up methods are far more taxing to produce, which may help explain why so few have used these methods. However, the large amount of potential error apparent in the top-down methods as described here may serve as motivation to lean toward bottom-up methods.

Missing research and the derivation of values. An important caveat is in order when speaking to the above estimates. In reviewing the literature, a troubling theme of missing research emerges. Program evaluations which account for costs of criminal justice exposure "plug in" the value of these costs based on estimates in the literature. However, some of the primary studies cited can not be found.

For example, studies such as Salome et al. 2003; French et al. 2002, French et al. 2000; and Rajkumar and French 1997 each "plugged in" costs to criminal justice agencies which they cite as coming from Rajkumar and French (1994). However, the research which generated these values was an unpublished working paper produced by the Research Triangle Institute (RTI), a private research agency. No electronic or paper copies could be identified through usual library-based methods. Likewise, the publishers of this report could not locate a copy (the head of research with the group has indicated that the study no longer exists and they have kept no copies of the document.[17] More specifically, the values people used in their research (the "plug in values") were taken from French and Rajkumar 1997, a later publication that listed the figures and referred to the earlier paper for a description of the methods. However, there is no known copy of the paper or description of the methods by which the values were generated (other than vague statements in Rajkumar and French 1997). Again, the problem here is that readers do not actually know what nuanced issues were associated with these estimates, even though other analysts could "plug in" their values if they chose.

Likewise, Crumpton et al. (2004) did not adequately report the values assigned to many key events in their analysis. Readers can not examine the estimates of court costs per case, the cost of booking, the cost of prosecution to judges, the cost of a hearing, etc.

[17] Unfortunately, the primary authors have not responded to requests for information on this report.

In fact, it was only with creative extrapolation of summary data and algebra that I was able to deduce their estimate of the cost of arrest to police; a feat not possible with their other cost elements. In short, we must take their word that their conclusions and estimates of the final cost and benefit associated with their recidivists are accurate. Thus, a second primary source of valuation estimates is un-reviewable.

In a more renowned study, Austin (1986) used "plug in" estimates taken from a study by Larson 1983. However, the Larson (1983) study was unpublished, no copies can be found, and no published description of the methods used exist other than some short description made by Austin 1986.[18] Thus, a third key source of "plug-in" values must be taken at face value, with little detailed information available on the sources of data, nuances of computation, or assumptions made to generate those values.

Aside from these issues surrounding computation of criminal justice costs, there is the remaining issue of the costs to victims. Again, this is a central issue to cost-benefit research, and represents the key split between different models for approaching cost-benefit research. There is controversy regarding (a) whether to consider victim losses (rather than just the criminal justice costs of crime), as well as (b) the appropriate estimation procedures and values attached to victim losses.

2. Computing Components: Valuating Victimization

Historically, "[t]he most significant (and controversial) portion of a criminal justice policy benefit-cost analysis is likely to be the cost of criminal victimization – in particular, the variation of intangible losses such as pain, suffering, and lost quality of life" (Cohen 2001, p. 23). Less controversy is focused on methods for deriving *tangible*

[18] Unfortunately, the primary author has not responded to requests for information on this report.

losses, although there is certainly variation in approaches endorsed. The more heated debate centers on identifying appropriate and accurate methods for attaching monetary estimates to abstract and intangible aspects of being victimized. How do you put a price on emotions or fear? And how do you measure the impact of crime on social factors like business disinvestment, or neighbors refusing to leave their homes after dark? These are real effects of crime, but controversy surrounds attempts to monetize them because many assume estimation is necessarily imprecise.

The controversy is less heated today than in past decades, largely because more rigorous and cross-referenced methodologies have been used to generate increasingly robust estimates of various tangible and intangible costs for victims and communities. Furthermore, the controversy seems less potent because there is a growing consensus that even imperfect estimates are better than no estimates at all. Indeed, even if some elements of crime costs remain outside of the econometric estimates, this is not necessarily a fatal flaw. Policymakers in particular seem to recognize that the neglect of some benefits or costs (those which are indefinable) does not mean that work can not proceed. Rather, it means policymakers must use information both from the cost-benefit data as well as substantive knowledge from other sources to make decisions. This is, of course, what most policymakers already do. If policy only required a list of facts and a common metric, then the state could simply program a computer to make policy decisions. This is clearly not the case. Instead, cost-benefit analysis is only one piece of information (although a powerful one) that should play a role in policy choice.

That being said, there has been exciting developments in the push to provide information to policymakers on topics that were once assumed out of the domain of

economists. Through creative and diverse methodological innovations, reasonable estimates of the monetary value of a variety of consequences of crime have been estimated.

As mentioned above, some aspects of estimating cost values for an event seem rather simple; other aspects of the estimate seem prohibitive. Several tools have been adapted from economists to facilitate estimation of costs associated with fundamental events in criminological analyses.

It is important to be clear on what it is economists are trying to do and why. They want to know what the value is that people place on each crime, just like other goods or behaviors, be it the value of beer, going to the movies, or having Microsoft stock. In the case of other market goods (i.e. the value of Microsoft stock) this is easy, because there is a directly observed market showing what people value the stock at (Aos et al. 2001). However, there is no such index for burglary or other victimization experiences. Thus, economists have to use other methods to elicit the value of something that is hidden from plain view, often referred to as "shadow pricing." The point is that economists are trying to determine one of two things: (1) either what a person would be "willing to pay" to avoid each crime; to determine a market value reflecting the demand of safety. Or, (2) what a person would be "willing to accept" in order to be victimized. These two approaches are analogous, but not identical. They may provide the same estimate of the value of victimization or different; the first reflecting an upper bound to the cost of crime, and the second a lower bound (Roman 2006, personal communication).

The more prominent methodologies include, "jury awards," "contingent valuation," "hedonic pricing / revealed preferences," as well as more direct methods such

as using NCVS survey data or the combination of injury data with estimates of costs to treat those injuries (i.e. public health or insurance estimates). Each method has strengths and limitations. Some only speak to tangible losses; some speak to the full range of costs endured by victims. The key is that by combining and cross-referencing information derived from each, increasingly reasonable estimates emerge. These methods are discussed in more detail below (see Table 2 for a summary of victim-valuation estimates across the published literature).

[Insert Table 2]

Victim Surveys

Victim surveys are most useful for estimating tangible victim losses; costs which easily convert to dollar-metrics. This does not, however, mean that measurement or accounting for their existence is in any sense simple or routine. For example, in the case of a robbery, the direct and tangible losses to victims may include the monetary loss, the cost of a hospitalization if injured in the process, and lost days at work while recovering from injuries or trauma. Much of this information can be captured through survey data, such as the NCVS. Victims are asked about property loss derived from recent victimization, injuries endured, and other direct and immediate cost information. The methodology underestimates losses for several reasons. The most important is the limited information victims have to draw upon. For example, the victims may not have an accurate idea of what their bills or lost wages will amount to, particularly because the short time-frame referenced in NCVS questions means expenses which mount over a life

time will not be known to victims.[19] Likewise, there are expenses victims may not be aware of. For example, Macmillan (2000) demonstrated that being victimized tends to decrease average earnings through the life course, primarily through disrupting educational and employment experiences. These events would likely take place after the time window of the NCVS has passed. To correct for this 'time frame' bias, as well as the concern that victims may not realize the costs they endured, Miller et al. (1996), and Cohen (1998) re-estimate the monetary costs to victims by combining NCVS data on injuries with public health research on lifetime costs associated with those injuries. Because of the care with which they corrected for these biases, the estimates generated by these authors are considered the most accurate in the field (see Cohen 2001, 2005).

Hedonic Pricing

In *'Hedonic Pricing'* studies, economists measure how people spend money to avoid events (such as crime). Examples include expenditures on car alarms, hand gun purchases, security systems, guard dogs, and house price fluctuations associated with crime rate changes. But they can also include other expenditures derived from analogous fields (i.e. willingness to pay for better physical health). By connecting the specific costs of these expenditures to specific goods/services people are making statements with regard to the collective value they place on the event in question. In as much as the value people attach to events is an interesting reference for the value of the goods/services in question,

[19] Correcting for some of these limitations, Miller et al. (1996), as well as Cohen (2005) draw upon medical research to estimate the actual long-term costs associated with specific injuries (i.e. a broken arm) so that estimates of medical costs, at least, are more realistic than the estimates derived from the short time-window of NCVS estimates.

this is a reasonable way to form a metric in estimating values of crime events as a whole (tangible and intangible victim costs combined).

For example, some have undertaken efforts to capture the value of crime rates as a function of the residual of housing market fluctuations (after partialing out other determinates of price). Even the casual reader will be taken aback by the gargantuan task this approach must entail – and indeed, there has only been moderate success at accumulating the data specificity and size to allow estimates to be captured. Hoen et al. (1987) went as far as to partial out factors relevant to home prices (such as characteristics of the house and the area) and factors relevant to wages in the area. Together, these data were used to calculate the marginal impact of crime rate changes (an index) on housing prices. Building on this approach, Bartley (2000) estimated a similar model while accounting for more specificity in crime types (rape, robbery, murder, etc.). Using estimates of the "willingness to pay" for a 10% reduction in crime across households, along with information on population and household sizes, Bartley estimated the cost of a rape to be between $6,875 - $72,192, a robbery between $2,239 – $4,094, and larceny between $8 - $89 per event. However, the ability to give more specific estimates was limited by lack of potentially omitted variables and lack of full accounting for the complexity of interactions between factors. Regardless, this line of inquiry represents one of the more creative approaches to capturing estimates of the social costs of crime. The method does so by attaching a monetary value to people's willingness to pay in order to avoid crime.

A related estimation process is the "compensating wage differentials" approach. This has most often been applied to estimating the value of a human life, and relies on

interpreting demands by workers for monetary compensation in exchange for doing risky jobs. For example, "[i]f persons demand $1,000 in compensation to take on a task with a 1-in-1,000 risk of death, the inference is that there are 1,000 persons who would collectively pay $1 million to avert this risk. Such an investment would be expected to save one of their lives" (Cohen 1998, p. 262). In an important expansion of the methodology to crime, Levitt and Venkatesh (1998) used accounting data provided by an organized drug gang in Chicago. The job of selling drugs was impressively dangerous, and the authors estimated the compensation demands by street dealers along with risk-of-death data. They found that the average drug dealer valued their lives at approximately $100,000. Importantly, this implies that the value of a life may vary across subgroups of people; the estimate is a function of subjective perception as well as other social forces.

Two caveats emerge in the literature utilizing this method. First, decisions are being made at the margin. The example above would not imply that individuals are willing to sell their lives for $1 million dollars. Rather, it is *risk* they are willing to take if given compensation, and the choice is a factor of the estimated probability multiplied by the value of the event. (However, the relative importance of the 'probability' component is conceptually non-linear.) Second, the bivariate estimates are inherently contaminated by other 'things' which impact wage demands. At the least, the amount of compensation demanded will be a function not only of the value associated with the risk, but by the supply and demand of labor, personal circumstances (need of money), opportunity for alternate jobs/choices, and quality of information. Regardless, the task for economists has been to capture the unbiased estimate of this value after partialing out these contaminates.

Limitations aside, there is impressive power in the use of "value-of-life" estimates that go far beyond estimating the value of a homicide (the most obvious use). For example, Cohen (1988) used supplemental homicide reports from the FBI to generate estimates for risk-of-death associated with a variety of offenses. With these in place, he drew upon the above estimates of the 'value of a life' to calculate the cost of each crime which derived from its unique risk of death (by multiplying the probability of death by the total number of those crimes – then weight it by the value of a life). The method has also been endorsed and used repeatedly by French and his colleagues in their evaluations of drug courts and other criminological interventions (i.e. French et al. 1997, French et al. 2000).

Jury Awards

In a second approach to valuation, Cohen (1988) included information derived from jury awards. The jury awards method makes intuitive sense: juries are directed to take into account the tangible and intangible costs to victims and translate that into a reasonable monetary value.[20] Jurists may not be experts, but they are people with the ability to empathize, to communicate what victimization is worth to citizens. More importantly, the combination of thousands of these "non-expert" data points allows patterns to emerge regarding the value that citizens estimate the harm of various crimes to victims. Although the value of a jury award is composed of several factors (i.e. the 'harm' which I am interested in here, as well as the wealth of the defendant, the composition of a jury, quality of evidence, etc.), multivariate techniques can be used to

[20] Most offenders do not have enough assets to make them worth suing civilly. However, there are cases in which a third party is sued (i.e. for not providing lighting or security in the area in which crime occurred) which can be used to estimate the value of the crime to victims.

partial out these contaminants and thus capture the independent value of 'harm.' In pursuing this method, Cohen (1988) matched information from the NCVS on average injuries by crime type with jury information on awards for those injuries (i.e. the value of a broken bone, a trunk injury, a lost finger, etc.). Likewise, Miller et al. (1996) used jury data and expanded the list of crimes beyond Cohen (1988). Unlike the earlier work by Cohen (1988), they also identified data in which juries were instructed to partial out their awards across categories of loss, such as "material loss" and "psychological suffering." Importantly, Cohen (1988) demonstrated that both "Jury Award" and "Willingness to Pay" estimation methods converge on similar estimates of the costs of crimes. Likewise, Miller et al. (1996) found that the "Jury Award" methods were consistent with other methodologies in generating estimates of the value of a life (see Cohen 2001).

Contingent Valuation

An alternative approach to valuation involves asking victims directly about their losses. Termed "Contingent Valuation," the method incorporates public surveys to generate estimates of victimization at the margins. It is, in fact, one of the more common and long-standing methods economists use across fields to estimate the cost of non-market goods and services. As Cohen 2001 notes, it is:

> A methodology developed in the environmental economic literature and has been used extensively to place dollar values on nonmarket goods such as improvements in air quality or saving endangered species. There have been, literally, hundreds of contingent valuation studies, metanalyses, and textbooks written on the subject (p. 38)

Despite the standard use in analogous fields, the methodology has barely begun to appear in the criminologist's tool kit; only two studies have incorporated the tool to date. In the first example, Cook and Ludwig (2000) surveyed citizens to determine how much they would be willing to pay for reductions in gun violence. In a similar study, Cohen et al. (2004) used a nationally representative sample of citizens over age 18 to investigate how much of a surplus amount of case (dictated by the study design) they would be willing to spend for programs which reduce crime (by a given percentage). They estimated the value (at the margins) for various crime types, and can illustrate how these estimates changed with or were robust to demographic differences of respondents. They estimated, for example, the 95% confidence interval value of a burglary to be $23,873 – $34,023 in tangible and intangible losses to victims (depending on the assumptions being made), and a rape at $209,817 – $354,987. This represented the willingness of citizens to pay to avoid the crime in question (at the margins).

3. Combining Components

The sections above focused on differences across the literature in what events should be considered a cost, and how those events are valuated. This section addresses the final way in which research has diverged: the *way* valuated components are combined. Five key areas of divergence emerge from the literature. These include, (1) deciding how much data-detail to collect, (2) dealing with 'multiple charges,' (3) conceptualizing the role of innocence in computations, (4) dealing with crimes which have not yet been valuated, and (5) what to do with 'victimless' crimes. Each is discussed below.

1. When to stop measuring and start assuming. A critical decision involves choosing which costs are assumed "fixed" across an event; when it is advisable to assign an average value to an event versus dividing it into further subgroups. For example, should we collect data on arrests only, assuming that an average number of those arrests will end in conviction and court time when calculating costs? Or, should we measure the actual rates of conviction? Likewise, should one assume costs are the same or different across crime types? Should we assume court costs are the same or different across dispositions? The same basic questions apply to police resources, court resources, sanction administration, and victim losses. Analysts must decide when to collect data and measure events, and when they can or should use an "average value" to attach to events. There has been substantial variation across the literature with respect to these decisions.

At the assumptive-extreme, Aos et al. (2001) used various state and national reports to estimate the number of arrests which would occur under various potential policy choices, then imputed the type of crimes that these arrests would represent, then imputed the number of arrests that would go to trial, then imputed the number of trials that would end in conviction, then imputed the average prison term that would occur, then imputed the average early release that would occur. No primary data or observations of actual events were made. Rather, a single figure is used to represent the "cost of crime" which is derived from a series of assumptions and imputations. In short, the only variable in their study was the estimated effect size of a program, which was multiplied by a numerical constant that was the 'average' cost of crime. The problem, of

course, is that programs are assumed to vary only in the proportion of arrests generated, but all other events are assumed identical (i.e. how many of those arrests will be for various crime types, result in convictions, lengths of prison terms, etc.). This is a problematic assumption.

Likewise, French et al. (2000) calculated the costs of recidivism by measuring the number of arrests for participants in various drug treatment programs. However, they did not collect actual data on court outcomes or sanctions. Rather, they used estimates derived from prior research to impute what would have happened among their sample, and then attach costs to those events. Further, they used a single number (a monetary constant) which represents the average criminal justice costs (police, courts, corrections), as well as the average victimization losses all bound up together into a single figure. In this case, the authors were assuming "crime costs" are static – unchanging across crime types or criminal justice decisions. The cost of a robbery arrest which endured trial and resulted in a prison term was assumed to cost the same as a shoplifting charge which was dropped at arraignment.

Moving a bit away from this extreme, Salome et al. (2003) evaluated a drug court and assumed that costs are dynamic *across* crime types, but constant *within*. In this case, the authors only collected information on arrests which occurred, which they separated by crime type categories. However, they did not collect any data on the actual court outcomes or sanction occurring amongst their data. Instead, they assumed that each crime type generated the same average cost (weighted by the average court outcomes and sanctions that the specific crime type generates in nationally representative data sets). Thus, there was a single figure which represented "the" cost of robbery, or rape, or

assault – no matter whether the arrest for their subject resulted in trial, prison, or was dropped. The assumption was made because the authors do not know the actual outcome of cases and thus impute the "average" court events based on national estimates of conviction rates and prison terms.

Zarkin et al. (2005) did not include estimates of police expenditures for recidivism, nor did they collect information on "crime types" amongst their recidivists. They did, however, include a more detailed estimation of court expenditures for recidivism processing than most studies. They did this by producing two figures which may represent court costs. The first was for cases which did not result in conviction – in which case they assumed one arraignment hearing occurred. Cases which later resulted in a conviction were assumed to have endured three hearings: an arraignment, a motions hearing, and a sentencing hearing. Using the figures from Czelen et al. (1978), they were computed a single figure to attach to recidivists, a figure which depended on whether a subject was reconvicted. Finally, they identified new prison terms for recidivists and attached a single per-day cost to prison sentences (derived as the average day-cost in the state of New York).

In short, the dominant theme across cost-benefit evaluations has been one of making simplifying assumptions about event occurrence, and thus about costs. At times, a single figure was used to represent all criminal justice expenses, or a series of "averages" were used which implied that court or victim losses were constant, rather than varying across crimes. At other times, cost to police were either ignored (Zarkin et al. 2005), assumed to be a single figure (Crumpton et al. 2004), or were derived from top-down studies which greatly inflated estimations (i.e. Salome et al. 2003, Aos et al. 2001,

Rajkumar and French 1997). Likewise, assumptions have been made which did not allow variation to occur in events which may truly be different, assumptions which may not be appropriate.

Why the assumptions? As can be imagined, the approach used often depends on the data available. The decision between what is measured and what is assumed constant appears less often to be made on theoretical grounds (i.e. actually believing that court costs are the same for dismissed charges and case which resulted in conviction). Rather, it more often represents the limitations a researcher reaches with respect to time or funding to go out and get data. These simplifying assumptions are usually accompanied by a foot note indicating that there was 'no data available' for an alternate approach, rather than a theoretical explanation as to why the choice was appropriate (i.e. Salome et al. 2003, Zarkin et al., 2005, Rajkumar and French 1997).

However, there are times in which these assumptions are expected and completely appropriate. For example, in the case of Cohen et al. (1994), Austin (1986), and Aos et al. (2001), the authors were interested in projecting the likely costs for hypothetical policy scenarios, rather than measuring post hoc events. Thus, the authors used "averages" in computing costs such as the average number of crimes expected, the average proportion that will be of each crime type, the average cost to courts (across all disposition types), and the average sentence costs (averaged across all sanctions).

Regardless, these "assumptive" methods are less useful to program evaluations in which cost-per-event is needed as a plug-in value (see Boardman et al. 2001). In the case of post-program evaluation, analysts do not want to simply measure the count of arrests and then assume there is an average number which are "robberies" and that there will be

an average number of those "robberies" which result in conviction, etc. Because we think the programs may generate a difference in behavior across the two groups in criminal behavior, we need to measure all realms in which that behavioral difference may emerge (i.e. number of arrests, type of crime, conviction rate, and sentence lengths). In each case, we want to measure the actual rates with which events occurred, because the alternate approach (assigning an "average value") can mask important differences which may occur between groups.

Again, the problem is that researchers may see differences in the rates of these post-arrest events for the treatment and control groups compared, even among people with the same charges. For example, there may be differences among groups in behavior which impact how cases are handled by the courts (be it the quality of acts alleged, the mitigation involved, innocence, etc.). A better assumption, then, may be that the difference in court costs may be *cumulative* across cost-inducing decision points. Thus, it is best to measure the actual difference in each event that may be related to behavior (i.e. the proportion of arrests, the proportion of arrests which are of each crime type, the court outcomes, etc.) for each individual.

To be clear, this does not mean that every possible event which can have a different value should be measured. Variables which are measured as "dynamic" costs should be theoretically linked to a program in question. Some variation is *not* substantively relevant and should not be included. For example, consider the recidivism event of "traffic violation." Should we assume a constant cost of 'police time' during traffic stops, or should we account for the actual officer's salary during each traffic stop our subject generated? Theoretically, we should assume a constant pay rate here, because

salary difference among officers engaging in the traffic stop is largely a function of luck, rather than a behavioral quality that may be due to the intervention in question. That is, the luck of having a twenty-year veteran at a high salary versus a first-year rooky with a low salary pulling a subject over is not something we want to give weight to – because it has more to do with luck than behavioral differences.[21] In short, there is not a rule that "the more the detail, the better the study."

2. Controlling offense or multiple charges? A second "combining components" issue in the literature has to do with accounting for charges among recidivists. All studies to date use the "controlling offense" with which a person is charged (the single most serious charge). Thus, a person who is charged with five counts of rape is valuated the same as a person with a single count of rape (all else equal). In this scenario, the multiple charges may derive from obtaining five separate victims, victimizing the same person five separate times, or a combination of the two. There is no literature to date which examines this issue, such as the linearity or cumulative nature of costs across multiple victims / single victim with multiple victimizations. However, it is plausible that the costs are higher among events which generated more charges.[22] At the least, then,

[21] Statistically, we should expect these "luck" differences in wage of officer to balance out to zero in large samples (i.e. as we approach an infinite number of traffic stops). However, most data sets are small (such as ours) containing only a few hundred subjects and a small number of traffic stops. Thus, the small count of events does not allow statistical processes of error-balance to occur and can result in unnecessary error. Using the average value here corrects for this.

[22] A likely reason for the omission (other than the difficulty of coding such data) is that there is often little information to differentiate the true multiple-victim / multiple-victimization events from instances of over-charging. Two events which look identical on paper (i.e. three charges of Robbery) may result from a single victim losing three pieces of property, or three people each losing their wallet – depending on the discretion of the charging officer.

evaluations may benefit from articulating differences in total charges before turning to analyses derived from a controlling offense rubric.[23]

3. The role of actual innocence. Some people are arrested who are not guilty of the charges in question. Most evaluations do not address this possibility explicitly, but do make implicit assumptions regardless. More specifically, most studies count the costs to police in making an arrest, the courts in holding hearings, the jails in maintaining detention during trial of *all* arrested subjects. If the person is found guilty, then the sanction is added in. If the person is acquitted or charges are dropped, then the cost is assumed to stop at that point in the justice system. Thus, most evaluations assume the subject's behavior generated a cost to the state which is represented by the valuation of these events, whether or not the person was wrongfully accused. This may be referred to as the "observed cost" of subjects. It seems unreasonable, however, to penalize a program for generating subjects who were wrongfully arrested; because that computational decision implies the program in some way failed when in fact it did not. Thus, in an ideal analysis, analysts would only count the costs generated by people who were actually guilty of the charges in question.

The challenge for program evaluation, then, is to identify which subjects are actually innocent versus actually guilty. This is far from simple when relying on official records, even with post hoc data on court outcomes. The problem is that we have little

[23] In addition, this may imply that future research would benefit from collecting data at the event level (i.e. police reports) in order to better estimate the nature of the act in question, the number of victims, as distinguished from simple 'over charging' of single criminal acts. To date, no study of this nature has been done, and the literature would benefit from evaluations which contain this level of event-detail, as well as valuation studies which explore the linearity and cumulative nature of crime-costs that involve multiple victims or multiple victimizations.

help in distinguishing legal from actual innocence. For example, a case could be 'nolle possessed' because of technical reasons or because a charged offender was found to possess a solid alibi. In fact, a person's charges may be dropped for a variety of reasons other than a lack of guilt (i.e. lack of evidence, technical problems with cases, unwillingness of juries to convict, etc.). Indeed, the presence of an acquittal, "STET," "Nolle Pros," "Dismissed" as well as "Not Guilty" verdict in court does not necessarily tell us about actual innocence. Likewise, it is also possible for factually innocent people to be found guilty both through trial and plea bargaining (see Forst 2004 for a review). In short, knowing what court outcomes occurred does not necessarily tell us about behavior.

This issue has not yet been addressed in the cost-benefit literature. However, a reasonable solution would be to run the analysis both ways, in order to establish the sensitivity of findings to varying assumptions. That is, compute the costs under the assumption that losses actually endured will be counted (i.e. police costs) regardless of potential innocence. The results can then be replicated under the alternative assumption that CJ processing costs are only "allowed" to count if the person was found guilty. Although the truth will lay somewhere in between (and neither estimate should be considered fully accurate), the two estimates then create a lower and upper bound of potential costs – a useful window to use when gauging program success (because 'truth' will be somewhere in between).

4. *What to do when there isn't a plug in value*?

The valuation literature has only addressed a limited number of crimes. This should not be surprising; there are hundreds of crimes specified across local, state, and

federal law. Two approaches emerge: either exclude the crime from the analyses (assuming the value is zero) or impute the value from analogous crimes. Both approaches have benefits and limitations.

One option is to place un-valuated crimes in analogous categories, assuming costs are about the same. For example, there are no economic studies detailing the cost of "resisting arrest." However, there are studies which estimate tangible and intangible costs of minor assault. It is plausible, then, to use minor assault as a "best guess" of the costs of resisting arrest. The problem arises when crime types do not have a defensible category to act as a proxy. For example, in the case of the early Perry Preschool evaluations, Barnet (1985) had no estimate of the cost of Rape, so combined this offense with "assault," which likely resulted in an undercounting of the cost of the crime (several additional examples can be found in Barnet 1985:47- 48).

The second approach has been to omit crimes which lack valuation research. If there is a general pattern in which one group is favored among the valuated charges, and the same direction of findings are present among the excluded items, then the omission is less troubling. It simply means, had there been better information on these other offenses, the findings would have been additionally reinforced. The problem, however, would emerge in the case that omitted crimes exhibited equal rates across the groups (if costly crimes) or proportions opposite the included crimes. This is not implausible. As an example, the Perry evaluation excluded: attempted robbery, possession of stolen car, attempted rape, sex offense, possession of stolen property, homicide, and all traffic offenses. They do not report which group was favored as far as these crimes are concerned, but they represent serious enough offenses that the cost-benefit results should

be interpreted cautiously considering their omission. Again, the exclusion of these costs is the same as assuming their value is zero, an assumption that is probably more wrong than assuming the cost of an analogous value which has been valuated.

Given the two options (omit or impute), it seems most reasonable to impute, particularly if the case is clear cut. However, it also seems reasonable to state these decisions explicitly so that readers can follow exactly which offenses where grouped together, gauge the impact of grouping decisions on findings, and test whether any imputation decisions made alter the findings in a substantive way (a sensitivity analysis).

5. What about victimless crimes?

Crimes which do not contain victims (i.e. prostitution, gambling, status offenses, and Drug arrests) represent a unique situation for valuation. As with other crimes, they involve direct losses to law enforcement. Yet there is considerable debate as to the costs they invoke to subjects involved in the offenses as well as the broader community. The debate is, perhaps, most well developed in terms of drug offenses.

Drug arrests may come as part of a routine traffic stop (fairly inexpensive) or as a result of a sting operation involving one or more cooperating agencies, diverse technologies, and substantial resources. In this case, the arrest would be expensive. There is no extant research which details costs across these diverse policing strategies. In addition, recidivism data does not usually give detail as to which source was used for a specific arrest. Indeed, most recidivism data does not necessarily differentiate well whether the arrest was for selling or personal use, the type of drug involved, etc.

Although most people agree that drug arrests cost criminal agencies money, there is a substantial debate regarding victim and social costs. Proponents of counting social costs note that drug use entails social burdens, including black market violence associated with regulating market transactions; violence or crime during use; or resulting from crimes committed in order to gain money to buy. Likewise, they argue drug abuse is associated with tremendous health risks and medical / hospitalization costs which are wholly, or in part, paid for by the community because users tend not to have health insurance. They also argue that drug abuse leads to unemployment or underemployment, as well as more injury on the job. Finally, some question whether parenting skills and episodes of child abuse may increase among users. These costs are highly variable across drug types, and user intensity. Coming from this perspective, several authors have attempted to estimate these various costs across drug abuse, and differentiate by type of drug being abused (i.e. see Rice 1989, MacCoun and Reuter, 2000).

Regarding productivity, Rice et al. (1989) used multivariate regression techniques to capture decrease in wages associated with drug abuse. They found an age interaction, indicating the average 18-24 year old who abuses substances lost $119 a year, the average 25-34 year old lost $833 per year, and those over age 35 earned approximately $4,336 per year less than non users. This value was estimated across heavy and non heavy users, and they argued that the value of "productivity represented the substantive value lost to the community (a market 'good' that would have been gained by the community had it been created). An additional loss of $578 per year is passed to the community from medical bills / hospitalizations in which the community picks up the bill (DOJ, 1992). Additionally, 3 – 14% of heavy drug users are killed annually (either due

to medical complication, overdose, suicide, or homicide). Cohen (1998) combined productivity data and risk of death to estimate that the average heavy user's risk of death laid between $1,485 and $6,690 per year. This is limited to heavy users only, and is not differentiated by age, drug type, etc.

Other costs involve the estimated average number of offenses drug abusers commit specifically to gain money for drug purchases (from self report data). For example, Rajkumar and French (1997) analyzed survey data obtained from 2,420 drug abusers arriving for treatment. They noted the self reported types and amounts of crime committed in the year preceding the interview and suggested this crime was a function of drug abuse, which they then attached a "cost" to, derived from Miller et al. (1996) and other sources on the tangible and intangible costs of crime.

In French et al. (2004), the authors argued that drug abusers are more likely to be both perpetrators of predatory crime, and victims of predatory crime. This paper was unique in that it expanded the conceptualization of drug abusers to consider them both as risks and at-risk; a novel and insightful approach. This perspective led the authors to estimate that the abuse of substances leads to considerably more cost to society than viewing abusers as "potential criminals" only. In fact, in many cases, the cost to drug abuse was more pronounced in terms of their risk of being victimized.

In another novel approach, Reuter used market prices of illegal drugs to estimate their cost to the public in terms of market violence. That is, he uses drug purchase prices which are "marked up" to account for the danger of dealing (approximately 50%), to represent the value of the product in terms of market violence. The price markup is a proxy for the willingness of drug dealers to sell drugs, representing their agreed upon

market value of the danger for selling. As such, these authors suggested using approximately half the street value of the drug in question to represent the social cost of black market violence represented in a drug sales event (or arrest).

Across these studies, authors were referring to the cost of a "statistical" drug offender; representing the average loss across all drug dealers or abusers. Some will cost more, some will cost less. And any one drug dealer or drug abuser in a particular sample may generate none of these costs at all. This is similar to other offense valuation, such as rape, robbery, and murder. In each of these cases, researchers know there will be variation in impacts, but are interested in the average value of losses for each offense; the statistical victim.

However, there are also critics of the valuation of victimless crime, and drug use in particular. First, these critics argue that the assertion that there are 'victim costs for drug use' is not justified; because the exchange is voluntary (there is no externality). If a person loses their life from drug dealing or drug use, it was a free decision and not a legitimate event to place a value on. One can not say, for example, that a drug deal costs the community in terms of the increased suffering of users (i.e. diminished life style, health outcomes, injury, or death that may result while high) because that user voluntarily engaged those risks. They were not imposed.

These arguments are probably the most common, and they also do the least damage in terms of serious discussion of cost-benefit tools; they are more semantic than practical. The value of a drug user's life is not zero, as they would imply. The fact that the drug abusers and dealers endure losses because of a choice seems like a red herring; the point is that damage is done which will be observed by the users as well as the

broader community due to secondary costs (i.e. picking up health care bills, stray bullets from drug dealer disputes, reduced quality of life for residents in drug-dealing areas, etc.).

A second, and more astute criticism, focuses on the assumption that drugs *cause* harm. The causal link between crime and drug use, for example, is hard to defend based on the extant research literature. To be sure, the correlation is present. However, there is little quality research establishing that drug use causes predatory crime (i.e. the research has not adequately dealt with issues of spuriousness, potential interaction effects with demographics, contexts, situations, and causal ordering). In addition, some find evidence that drug use has benefits for communities. MacCoun and Reuter (2000) noted the average drug dealer in large cities invoked approximately $15,000 in financial transfers from affluent areas into poor communities. As drug dealers sometimes spend a substantial amount of profits derived from suburban dwellers on social welfare (i.e. feeding local children, paying rent for unemployed seniors in their neighborhood, buying school supplies for children, etc.) it is difficult to argue that drug dealing only involves social costs.

Finally, a third argument criticizing the cost-of-drug literature centers on the use of intermediate events (like drug use) as proxies for drug losses. If an analyst attaches a monetary value of "theft" that the average drug abusers commits in order to purchase drugs, then she is making a bold assumption regarding our particular drug dealer or drug user. The critics are skeptical of the assumption and argue that analysts should instead count whether or not the individual has been arrested for theft, rather than making assumptions. As noted above, proponents tend to counter this argument by noting that

most of valuation economics focuses on determining the cost of a statistical event, rather than the actual observation in the case at hand. It is the pursuit and use of "average costs associated with events" that forms the foundation of cost-benefit valuation, because that is what we are actually interested in: the generalized value of the event; the average cost associated with an event.

The debate is far from settled. Given the tentative and polarized nature of drug-cost estimates, along with the lack of information in most official record data regarding the type and amount of drugs involved (as well as the intention of selling versus using those drugs) analysts should be hesitant in their use; including sensitivity analyses if they are used.

Current Research Questions:

The current study is implemented in order to answer several specific questions. In answering each, I attempt to improve upon previous gaps in the literature which were described above. These questions include:

1) What are the costs of operating each program? What are the post-program costs generated by each program?

2) Which policy option is the most cost-beneficial?

3) How robust are these cost-benefit findings to changes in coding schemes (i.e. changes in cost-benefit models, valuation estimates chosen, or combining of components)?

CHAPTER 2: METHODOLOGY

In this section, methodology and procedures are outlined. Again, the goal of this project was to evaluate the cost and benefits of the Toulson Correctional Boot Camp MAP program, compared to a policy option of operating the MAP program at a traditional prison instead. Description focuses on how data were collected, how subjects were chosen, and how variables were coded. The section also describes the facilities used in the evaluation and the data bases from which information was derived.

Facilities

Two facilities were used in this evaluation. The treatment group was kept at the Toulson Correctional Boot Camp (BC), a prison facility located in Jessup, MD. The facility was designed to house approximately 450 inmates, and rarely exceeds the designed capacity. All inmates housed were low-risk inmates who qualified for the boot camp, meaning they were first-time adult inmates who had no violent convictions.[24] The complex consisted of several single-story buildings serving as a school, a cafeteria, and housing / administrative areas. The grounds consisted of obstacle courses, a running track, and well groomed lawns with flower beds. Additionally, there were trailers in one section of the complex that operated as "privileged housing" for platoons which were near graduation and had earned the extra independence and privacy. Program Inmates lived in dormitory style housing, similar to therapeutic communities. Their time was highly structured no matter where they were or what time of day it was. Their schedule began early in the morning, with small amounts of time allotted for hygiene tasks (such

[24] The facility also "rents out" its space in its "staging wing" to low risk inmates who are over-flows from other institutions. These inmates are often older inmates and do not come in contact with program inmates.

as showers, cleaning bunks, eating) and all were conducted in military fashion (i.e. bunks were made to military specifications; clothes were arranged according to military fashion, and inmates ate meals in silence while sitting with military posture). Likewise, inmates were required to interact with each other and staff in a militarized style, including the use of "sir/ma'am" when speaking to staff, standing at attention, and reporting rules of the facility on demand. The staff used summary punishment for rule infractions (i.e. doing pushups, or some other form of exercise for forgetting a rule, speaking out of turn, or making other mistakes). The camp emphasized therapeutic interventions (i.e. scheduled time spent in education, life skills, and substance abuse therapy) as well as vocational training (i.e. work crews) and physical fitness. Not only were inmates required to participate in programming, but a drill instructor sat through all classes with the platoon and "graded" them on the attentiveness, effort, and manners when in classrooms. Finally, the facility was noteworthy for its lack of privileges, such as highly restricted access to visits, phones, no television, no music or radios, and very limited personal property (i.e. no personal clothes, hats, decorations, etc.).

The comparison site was the Metropolitan Transition Center (MTC), a traditional prison facility. It was built in the late 1800s and is the oldest operating facility in the United States, serving as a secured pre-release center. It is located in Baltimore City and was built to house approximately 1200 inmates, although it averages between 1400-1500 inmates per year. The inmates were all serving the last two years on their term when transferred to the prison. This means the general population could have been very low risk inmates (never had more than two years to start with) or high risk inmates who had served time elsewhere but were preparing to reenter the community (i.e. inmates could

have been serious and violent offenders who were nearing the last two years of a very long sentence). The facility did not systematically separate inmates into housing units according to risk – they were housed largely at random and based on bed space availability. Research subjects sent to this facility were prioritized for access to the same treatment services as provided at the boot camp (education, substance abuse treatment, and life skills training). However, the facility demonstrated less efficiency in getting inmates into programs, or getting them into the programs quickly (implying less treatment dosage). Although inmates were required to follow basic rules such as speaking respectfully to guards, to avoid contraband, to avoid violence, etc., there were fewer demands placed on inmates at this facility. There was no rule that inmates must be able to recite facility rules on demand while standing at attention, and there was no rule against physical contact or yelling as in the boot camp. Likewise, there was routine access to items that are not allowed at the boot camp. These include personal property and clothing, televisions (communal and personal), radios, microwaves, recreation such as basketball courts and weight lifting equipment, and, most importantly, free time. Table 3 below describes facility characteristics derived from annual performance measure data submitted to the state legislature by each facility.[25]

[Insert Table 3]

[25] The data source for these performance measure data are described below.

Participants

All inmates who were sent by the court to the boot camp program for a 'six-and-out' term between 2001 and 2003 were used for the evaluation, N=234.[26] All inmates were included in the study and all inmates were randomly assigned to serve their MAP contract term at the Toulson Correctional Boot Camp or the Control site. Although inmates could chose not to participate in the self report sections of this study, their facility assignment was determined according to the randomization process without choice (assignment decisions are at the discretion of DOC, inmates are never allowed discretion in this regard).

Out of the 234 participants in the study, eight were excluded from analysis. One was excluded because he was deported upon completion of the MAP contract. The remaining seven inmates were not free the minimum 12 months required in order to contribute to the analysis of post-program behavior. There were no differences in the rate of excluded inmates between the two facilities, and no significant difference between excluded and included inmates in terms of criminal history or demographics (data not shown). This results in a final sample size of 226.[27]

---

[26] The Parole Department determined eligibility for the MAP program. An individual's offense and criminal history score determined whether he/she was eligible for the boot camp program as a Part 1A offender. Specifically, inmates could not have been convicted of a violent offense currently or in the past (although they may have been charged with them). All potential MAP inmates were housed in an isolated section of the boot camp designated to house incoming inmates at TBC. As their program eligibility was being evaluated, they were kept separated from the program inmates and were not involved with the boot camp atmosphere in any meaningful way (i.e. they were not required to say 'sir' or act in a manner required of program inmates. Likewise, they had no contact with program inmates).

[27] More specifically, these seven inmates were excluded because the limited follow up period would make them look like "non-recidivists" whether or not they were in truth recidivists. The lack of follow up time meant that both facilities might look slightly more effective than they were, and slightly more similar than they may have been. Because I am unsure of the validity of their "post-program" value, these seven inmates were considered to be conceptually similar to "missing values" on post-release follow up costs associated with recidivism. Likewise, the single deported inmate was held for extradition upon release, which implies a cost (i.e. deportation hearing and detention) but not a cost which can be attributable to the

Random Assignment

The researchers arrived at the boot camp one week before each new platoon was scheduled to begin the boot camp program. Upon arrival, the researchers were given a list of the inmates scheduled to begin the program. Through the use of a random number generator, the researchers determined whether inmates were selected for the boot camp or comparison facility. The random assignment decisions were final. Neither the research team nor any correctional employee could change the decision once made.[28] The process resulted in 105 inmates sent to the boot camp, and 121 sent to the comparison site (the differences in sample size were no different than expected by chance alone). See Table 4 below for a description of group the sample and eligibility.

[Insert Table 4]

Data Bases

1. Inmate Surveys

The research team administered a survey to all subjects several days before they were to begin serving their MAP sentence (all subjects were housed at the boot camp staging area). Although participation in the survey was voluntary, survey completion rates were exceptionally high (223 out of 226; a 98.7% participation rate). Questions

---

programs in question. Again, his post-release costs would artificially appear to be zero with respect to recidivism due to being located in another nation, which would have entered a small amount of error into our computations.

[28] In one case, an assignment error occurred. A new administrator at the boot camp misunderstood procedures and sent 6 inmates assigned to the boot camp to serve their term at the control site instead. These 6 were kept in the data set and listed as boot camp participants regardless. This may create a slight bias in which the effect of the true boot camp effect appears less than it would have been.

and response options were read to the groups by research staff in an isolated room. The survey took approximately 45 minutes to complete. The survey data were used to generate descriptive statistics which were cross validated with official data from police and parole records when possible (i.e. age, race).

2. Department of Corrections Data Base (OBSIS I)

The Maryland Department of Corrections maintains a data base which tracks inmate characteristics and events while detained. The data base has rich information regarding movement and locations within detention centers, including dates of location change and reasons for changes. The data was primarily used to decipher start and release dates of inmates in the competing programs, as described in the "variables" section below.

3. Parole Data Base (OBSIS II)

The Maryland Department of Parole and Probation maintains a data base which tracks offenders while on parole (which includes all subjects in this study). The data covers demographics, sentence information, as well as the tracking of key events (such as arrests and warrants for arrest). In the case of arrest events, the date of offenses and charges are also listed. The database was primarily used for cross referencing official police records of recidivism.

4. Criminal Justice Information System (CJIS)

The state of Maryland maintains a central repository of information on arrests and court outcomes. The data base contains a list of charges filed, the jurisdiction and dates of charges filed, and court events (bail, disposition, and a sentence information). This is a key source used to track both past arrest history of subjects, as well as recidivism events and timing of recidivism.

5. Legislative Budgets and Performance Measure Reports

The state of Maryland publishes budget data for the Department of Public Safety annually. Within these budgets are individual lines of expenditures and revenues (i.e. commissary and inmate labor profits) associated with the two facilities in the study. The same reports supply information referred to as "performance measures," which include the average population level of the facilities, as well as staffing levels and critical event counts (such as escapes or injuries), as described in Table 3 above. See Table 5, 6, and 7 for a list of expenditures and income generated by the two facilities for each of the three years covered in this evaluation.[29]

[Insert Tables 5, 6, and 7]

Variables

1. In-program: Time Served

The MAP program dictated a 6-month term for all subjects (183 days). However, the "days served" could change under certain conditions. Inmates could be released early

[29] All monetary values converted to 2005-dollars. Conversions to a common dollar-metric were completed using techniques recommended by the Federal Reserve. For more information, see http://woodrow.mpls.frb.fed.us/research/data/us/calc/

due to (1) a successful appeal, or (2) judicial intervention ordering a release. Inmates could have their terms extended if they "failed out" of the program. Inmates would fail-out if they were found guilty of a major infraction (such as assaulting another inmate or staff), or refused to participate in programming (such as education, life skills, and substance abuse treatment – each of which was mandated for MAP contract holders). If an inmate "failed out," one of three sentence increases occurred. This included (1) a 30-day addition, (2) a 60-day addition, or (3) revocation of the MAP contract. In the later option, the inmate lost all credit for "good time" earned and reverted to their original sentence (minus actual days served). To estimate the actual time served, DOC data (OBSIS I) was searched to identify exit dates from facilities, and exit reasons. By comparing this actual-exit to expected-exit dates derived from MAP contracts, difference in expected and actual release times were identified, as well as reasons for changes in release dates.

2. In-Program: Facility Costs

As described above, budget data were obtained for each facility through published audits maintained by the state legislature. After converting dollar values to a common metric (2005-dollars), the three years worth of data were averaged in order to generate the mean annual expenditures for each facility. This information was combined with average daily population rates to generate a per-inmate cost as described in the results section below.

It's important to reiterate here that there are important conceptual and practical differences between the uses of marginal versus fixed day costs of inmates. The fixed

cost approach I use is a clean and easy to follow representation of cost the state realizes. By differencing the costs at each facility, the analysis gets an easy to understand and useful proxy of the difference in costs each program generated. However, many would argue that the preferable measure is the 'marginal' costs of a facility. The reason is that we can not be sure how the distribution of fixed versus marginal costs at two places being compared will work out, whether the average cost will accurately reflect the magnitude of the difference in how *budgets* will actually change over time.

For example, suppose we have two facilities. One costs an average of $60 for an inmate per-day, the other costs $80. The difference would be $20. However, suppose that have the costs are dynamic and half are fixed at each place. In the case we would say that the first facility has a marginal day-cost of $30 per inmate, and the second has a marginal cost of $40. The difference here is now only $10 per inmate per day. Which is right? It depends on the question. The marginal cost is often preferred in program evaluation, as it reflects real-time costs (changes in potential budgets). However, the two figures do often agree with respect to direction, and can often be similar in terms of magnitude. In addition, the marginal cost is often a difficult figure to estimate which means that its use may come at some analytic cost, at least in some situations.[30]

[30] For example, in this research, I attempted to calculate both average and marginal costs of the two facilities, in order to re-compute all the tests using both methods. This would allow a transparency that may be useful to policy makers, and serve as an exploration of sensitivity. To do this, I collected budgetary data and average inmate populations over 5 years for each facility. I computed the average change in population between each year (i.e. 2000 to 2001) and the resulting change in budget (i.e. 2000 to 2001). Repeating this for each year, I was able to come up with a series of figures that, substantively, reflected the change in budget per 'X' increase or decrease in population. By converting to a constant 'year-dollar' and carrying out the implied division, these differences could then be used to represent the marginal costs of adding just one more inmate to each facility. My goal was to average out these marginal costs over a 5 year period to account for any shocks. However, the exercise was hampered by the realization that the budgets did not track population changes at all. Not in magnitude or direction. Some years would have lower budget and more inmates. Some years would have no change in inmates and high jumps in budget. Thus, it was clear that there were omitted variables and the simply regression I was applying was

Although we know this proxy is composed of fixed and dynamic parts (i.e. fixed costs of operation that wouldn't change by adding another inmate and the costs that would), the average day-cost is still useful in many instances. I use it here as a clean and clearly measured proxy of the difference in costs.

3. Post-Program: Time at Risk

The inmates entered the study in monthly cohorts; one cohort per month between 2001 and 2003. This was dictated by the structure of the program being evaluated, which started new "platoons" as cohorts once per month. The recidivism data was downloaded on a single date in November 2005. Thus, inmates varied in the amount of time they were at risk to re-offend between (a) release and (b) November 2005.

The cost-benefit analysis is a group level comparison, and thus the key question mergers: did the two groups have equal amounts of time (in the aggregate) to generate costs in the form of recidivism? If the groups had different times at risk, then any model comparing sums of their costs would have to account for this. In contrast, if time is equal, then we need only employ simply summing measures, converted to per-person metrics.

Indeed, the two groups had equal time at risk on the average. The boot camp inmates were "at risk" an average of 808 days, whereas the control inmates were "at risk" an average of 803 days (the averages are not significantly different than expected by chance).

---

mispecified. Not having additional information on possible covariates, I then chose not to pursue the parametric estimates of marginal costs and instead chose to use the more clearly measured fixed costs.

4. Post Program: Recidivism Events

The analysis in this paper restricts consideration to new crime events; substantive criminal behavior rather than technical parole violations. Although data was obtained on technical parole violations occurring, they were rarely associated with any sanction other than a hearing in which the subject was continued on parole. Unfortunately, no budgetary data was available to price these events.

The recidivism data were downloaded in November 2005, and all recidivism data represent the state of criminal behavior (or lack thereof) as of this date. Two data sources were used to derive recidivism information. The first was the Criminal Justice Information System (CJIS).[31] The second was the Parole and Probation data base (OBSIS II). The two data bases were cross referenced in order to identify any arrests which were listed on one, but not the other. A complete list of offenses, dates of offense, charge names, and court outcomes was created. Both data sets and the coding procedure are described below.

A. Police recidivism data (CJIS)

All adult arrests recorded by police officers in the state of Maryland are entered into the CJIS system. This system is also linked with the courts, and contains info on the date of arrest, charge names, and court outcomes (i.e. dispositions and sentences if applicable). Complete CJIS data were obtained for all subjects (i.e. for their lifetime). All

[31] I used information from parole data to identify subjects who moved to different states. Parole officials in these other states were contacted to request rap sheets on each subject be run and this data transferred to the research team.

arrests which occurred before release from confinement were removed.[32] This resulted in a list of all charges brought against the samples 226 subjects since release. In the CJIS data set, a total of 263 recidivism events were identified, all of which were for new crimes (none were parole violations). The number of events was influenced by offenders who generated more than one arrest; the most active having as many as 8 separate arrests in the follow up period.

B. Parole recidivism data (OBSIS II)

The parole data base listed all warrants issued as a result of a "new crime" arrest. Parole officers had discretion as to whether a warrant was issued once they were informed of an arrest, and they did not always issue a warrant. Their informal policy was to wait until charges had been both reported and substantiated by the courts before issuing a "new crime" warrant. For some officers, this may have meant a probable cause or arraignment hearing has passed with sufficient cause given for continuing the prosecution. Others may only have issued a new-crime warrant after a conviction, but not if the charges result in dropped charges or an acquittal.

It is important to recall, of course, that some subjects could have finished their parole term before a new-crime arrest occurred. In these cases, the offense would not be included in the parole records. Thus, far fewer "new crime" warrants are expected to appear in the parole data base, relative to the police data base (CJIS). However, all

[32] Data referring to arrests before program entry were used to compute variables referring to official criminal history, such as the number of arrests and the number of convictions. As described below, I identified 8 cases in which the CJIS system had no history whatsoever to the individuals; they were missing cases. In these cases, I contacted the subject's parole officer and had them research the individuals conduct since release and report all contact they had with police (i.e. arrests) since release to supplement official records from CJIS.

charges which did occur in the OBSIS II (parole) data base were expected to also be found in the CJIS records (police).

In all, 138 new-crime warrants were listed in the parole data base (between program exit and the end of parole or the censor date). The parole data base, however, had "double counting" of arrests, which came from two sources. First, it included multiple rows for the same subject (at times) because more than one active period of parole or probation may exist for the same person. In the event of a single arrest event, two or more rows in the data base may have listed that same event as a unique arrest (i.e. each officer monitoring a case hears about the charge and lists it as a new-crime warrant). Second, some arrest-events involved numerous charges; more than the "four" there is room for in the parole data base for each arrest event. Rather than picking the four most serious charges, some officers elected to enter the data as if there was a second unique arrest on the same day, and then continue the list of charges.

In total, there were 26 events listed which were double-counts over 13 unique cases. This means there were 122 unique arrests events identified in the parole and probation data base. These 122 arrests were spread across 100 individuals. Of these, 79 had a single new-crime warrant, 20 individuals had two unique new-crime warrants each, and 1 subject had three unique new-crime warrants issued.

C. Cross checking police and parole data: Coding for inclusion

The two recidivism data sources (Police and Parole) should have had some, but not perfect overlap. The Police data (CJIS) should have listed every arrest which occurred, whereas the Parole data (OBSIS II) should have listed the subset of those

arrests for which the subject was on parole during the arrest and the parole officers decided to pursue a "new crime" warrant.[33] The two data bases were compared in order to cross check the accuracy of the police data. That is, every unique arrest listed in the parole data (n=122) was compared to the police data base listings for the same person (all arrests in the police data base for that person) to ensure that all were represented in the police files. The police data contained unique identifiers for each arrest event (a tracking number as well as a case number), but the parole warrant data did not. Thus, the parole files had to be reconciled by hand with the police data base. To accomplish this, the date of arrest, name of subject, and name of charges were compared between sets in order to determine whether an arrest listed in the parole data was indeed replicated in the police data set. Decision rules turned out to be obvious, as there were no grey area cases in which the choice of events matching were unclear.

Again, the police data should contain every arrest listed in the parole file. However, our cross-reference methods indicated there was substantial missing data from the police records (CJIS). Out of the 122 unique arrests in the parole file, only 85 (69.7%) were in the Police data. This means that a total of 37 (30.3%) of the parole new-arrest warrants were not recorded in the central police data base. Part of the explanation is that the parole agents listed arrests derived from traffic events. That is, actual arrests and bookings (not traffic tickets). In contrast, police records from CJIS do not list traffic offenses, even if they end in arrest. However, only 7 of the 37 "missing" arrests were

[33] One avenue by which a person may appear for a criminal charge, but without an arrest/booking occurring (which would mean the charge would not be in CJIS). This would be in the event of an offense for which the victim(s) swear-out a warrant, and an alleged offender appears in court to address that warrant. For example, in the case of domestic violence, a warrant may be issued and a court case may ensue – even if an actual arrest by police never occurs. However, this is (a) rare, and (b) does not seem to apply to the offenses which appear in Parole data, but not in Police data.

traffic arrests. The remainder were for a variety of offenses, many of which were serious violent offenses.

In pursuing this issue with Parole and Probation officials, the department suggested there may have been problems in which some police officers (and some police departments) did not report information to CJIS efficiently. Evidence to this effect is also found in the fact that 8 of our subjects were identified by CJIS as having *never* been arrested in their lives. However, these subjects were on parole (they are in the parole data base as active cases) and had been to prison (which is how they came to our study). Thus, they had been arrested at least once in their lives. Regardless, they could not be matched in CJIS with any known offenders using their State ID (SID), their FBI number, their social security number, their Department of Corrections Number, or their name/date of birth. This implied was indeed some room for error in the CJIS system.

Because of the large error rate in the police records, recidivism data was coded for inclusion from the two sources. This means that an arrest was assumed to have occurred if either one of the data bases listed an arrest. Overlapping arrest events were identified in both data bases to ensure there was no double counting (and, as described above, double counting within the parole data base was also corrected). Coding for inclusion, the data showed 144 recidivists, representing 300 unique arrest events and 905 charges. These arrest events are compared across facilities in the results chapter below.

5. Recidivism Cost: Making an Arrest

In certain cases, these subjects generated charges which were not in any specific valuation literature. For example, there have been no specific studies with offenses such

as "vagrancy" or "resisting arrest." In these cases, offenses were grouped along substantive grounds with a charge that had been valuated. For example, "resisting arrest" is assumed to be similar to non-injurious assault. Likewise, "vagrancy" (unevaluated) is substantively similar to "public disorder" (valuated). The decision of grouping cases was made based on the substantive similarity in the offenses and likely police response and resources.[34] For minor offenses which were not similar to other valuated categories, a minimal value of police resource used was assumed (one hour of one MD police officer's time, calculated at an entrance pay scale for officers: $20.25).[35] See Table 8 below for a list of offenses grouped together in valuated categories.

[Insert Table 8]

The monetary cost of making an arrest was derived from a variety of sources. Table 9 lists the values assigned to each arrest type. In computing the "best estimates," emphasis was given to bottom-up studies from the published literature, and those studies which engaged the most current, thorough, and generalizable methodologies (i.e. Miller et al. 1996, and Cohen et al. 1994).

[34] The alternate approach would be to exclude these charges from the analysis. However, it's important to note that this approach is still making an assumption about value in this case. Instead of assuming the value is similar for an analogous offense, the approach would be assuming the value is "zero." Although there is likely a small amount of error in assuming that "resisting arrest" uses approximately the same police time as "minor assault," the amount of error involved is likely far smaller than would be found in a method which assumed the value of police time was 'zero.'

[35] According to legislative reports, the average starting pay for a starting police officer was $17.00 per hour, + 25% fringe, which comes to $20.25 per hour. The estimate uses the lowest possible pay level in order to create a conservative bias against finding a boot camp effect. That is, because the boot camp generated fewer crimes over all, assumptions which result in low-estimates for crime costs result in making it harder for the boot camp to demonstrate benefits. However, if the boot camp still demonstrates benefits despite the uneven playing field (relative to the control site) there may be more confidence in the robustness of the benefits when making policy decisions.

[Insert Table 9]

6. Recidivism Cost: Court

A similar methodology to Zarkin et al. (2005) was employed in order to estimate court costs. Dismissed cases were assumed to have only one hearing (an arraignment).[36] However, *nolle prossed* or stetted cases were assumed to have two hearings (an arraignment and a minor hearing). Any case which resulted in a finding of "not guilty" was assumed to have three hearings (all but the sentencing hearing). Finally, a conviction was assumed to have one of each hearing type, for a total of four hearings. Thus, this coding is similar to Zarkin et al (2005), except this paper included more court outcomes and thus more variation in costs was allowed.[37] As shown in Table 10, these decisions brought the cost of a dismissed case to $653, a case in which prosecution was ceased (nolle pros or STET) came to $1,213, a 'not guilty' finding was assumed to cost $2,854, and a conviction was assumed to cost $3,763 in court related costs. Again, these were "bottom-up" estimates derived from consideration of facility resources, as well as salary data of all individuals involved in the criminal prosecution (judge, prosecutor, clerk, and security).[38]

---

[36] Importantly, this assumption creates a slight bias favoring the control site. That is, because 15% of the control site court cases resulted in a dismissed finding, but only 9% of the boot camps, the decision to count this category as less costly than in Zarkin et al. (2005) means that this papers approach attached relatively less weight to the control site than their method would have.

[37] To be clear, the substantive meaning of a "dismissal" versus STET or *nolle pros* are different, and the "dismissal" implies a more clear cut and easy decision in the dropping of charges. Each usually occurs at earlier stages of the court process or during the arraignment hearing.

[38] Unfortunately, data pertaining to the use of public defenders was not made available to the researchers. Although a reasonable approach would be to include the same number of hours for public defenders as prosecutors, a significant problem remains: we don't know how often a public defender was used. It may be often or rare on the average across subjects, and the rate may differ across groups. Therefore, any estimate of the use of PD services would be tenuous.

[Insert Table 10]

It is likely there were often more hearings than those assumed above. However, the "minimal hearings" assumption had two benefits. First, it created a conservative bias. For example, in as much as the boot camp generated fewer convictions, each conviction was likely under-weighted and the boot camps benefit were muted slightly by this method. This again created an uneven playing field making it artificially harder for the boot camp to demonstrate it was cost beneficial.

The second benefit being that weight was not given to the number of hearings which actually occurred, because the 'count' of hearings could have been a function of forces other than "behavioral" differences resulting from program assignment. For example, a greater number of hearings may have reflected a more complicated case, or a defendant who demanded a jury trial and refused to acquiesce to prosecutors. In the latter case (trial demands) the data would show a higher cost to the courts, but not because the behavior of the defendant was necessarily worse. In fact, the trial demand may be due to a greater rate of actual-innocence among one group. In short, there are confounding factors which may drive a case to have a greater number of hearings, aside from offense severity. Thus the Zarkin et al. (2005) method was chosen instead.

8. Recidivism Cost: Sanctions

Through the CJIS system and Public Safety records (OBSIS I and OBSIS II), all sentences resulting from new crime recidivism among the sample were identified.[39] This

[39] As noted above, a group of recidivism cases were identified in the parole data base, but were not found in CJIS. A group of researchers at the Department of Public Safety researched this group of cases by hand in

included information on the total sentence length ordered, as well as the minimal time ordered (total minus suspended time). As noted in the literature review above, there is reason to believe that specifying prison costs according to security level of recidivists may be beneficial. There is, of course, no way to predict exactly where each inmate will be housed – or which combinations of prisons the inmate will be housed in – over the duration of most sentences. However, there were general rules of thumb derived from classification policy in Maryland which were used to distinguish inmates who would likely serve their term in four different settings: Jail/ Pre-release, Minimum security prison, Medium security prison, and Maximum security prison. Sentences were classified by the following criteria:[40]

1. Jail / Pre-Release assumed for sentences less than one year
2. Minimum security assumed for sentences which were between 1 – 2 years in length.
3. Medium security assumed for all sentences that were greater than 2 years in length
4. Maximum security assumed for all sentences which were Homicide or Sexual Assault (regardless of term length).

---

order to locate the physical records of court events. These data were added into our CJIS data base and then used for the computations described.

[40] The state of Maryland uses an objective tool to classify inmates to security levels. Items include time-to-serve, offense severity, and other factors related to escape or rule violation risks (i.e. past history of problems while detained, past escapes, detainer for serious offense filed and number of prior incarcerations, and past convictions for violence). The inmates in this sample were substantively similar across all risk factors, as these were related to their eligibility for the boot camp program in the first place. For example, all subjects have no history of past convictions of violence. Thus, the only substantive criterion that varies across recidivists and which will impact their placement is (a) sentence length, and (b) offense severity. In terms of general practices for placement decisions in MD, the sentence length is the driving factor for classification, other than extreme cases of offense severity (i.e. a homicide or rape conviction may often result in an exception such that an inmate is classified to begin in a maximum setting). Although the classification scheme used to code these data is designed to be a lose approximation of the classification process, it should be a close approximation for the average recidivist in this sample.

To estimate the difference in cost, the most recent published budget data by the state of Maryland (the 2005 budget) and the net cost as well as average population levels of all potential prisons in Maryland were used. They were separated into the four substantive categories listed above, and the average (net) day-cost per inmate was computed within categories. The computation was limited to prisons for which these inmates would qualify (i.e. Women's institutions, super-max, and diagnostic facilities were eliminated from the estimates of costs). The data imply the cost of incarceration does differ substantively across these criteria, as shown in Table 11. These values were attached to the conviction information in later sections in order to "weight" the cost of incarceration by the resources differing sentence types will draw upon when administered.

[Insert Table 11]

9. Recidivism Cost: Victimization

As noted above, the literature has distinguished two categories of crime costs to victims. The first relies only on directly measurable losses, including lost property, lost time at work, and medical losses. The second involves econometrically derived estimates of pain and suffering of victims (the intangible losses) as well as the losses to communities or secondary victims (i.e. the impact of a homicide on the family of the victim in terms of lost wages, psychiatric visits, and pain and suffering). The literature's estimates were used as plug in values for the victim costs of crime (see Table 12 below). The majority of these 'best' estimates were derived from elements of Miller et al. (1996).

However, in the case of crimes which were not covered by Miller and colleagues, the plug in values from other sources were used. Again, priority was given based on the recentness of data, thoroughness of sources, and bottom-up methods.

[Insert Table 12]

CHAPTER 3: RESULTS

This analysis follows the same general framework of other cost-benefit program evaluations. Specifically, it progresses through three steps, including (1) computation of differences in program costs, and (2) a computation of differences in post-release costs (utilizing the tax-payer model, the direct costs model, and then a full cost-benefit model). Finally, (3) summary measures are presented to illustrate difference in costs and benefits generated by each of the three models used. Within each "step," sensitivity analyses are presented in order to illustrate the resilience of findings to varying assumptions.

Sample Description

Comparisons of the 226 offenders in the sample are shown in Table 13. Most were Black, young, and from Baltimore City serving time for a drug offense. On average, they had had more than 5 previous arrests and 2.5 prior convictions. Although there were no significant differences between groups with respect to demographics or personal characteristics, there was one important difference with respect to post-release environment. That is, significantly fewer control subjects had been assigned to intensive parole (ISP) after exiting their facility.

[Insert Table 13]

Table 14 provides additional descriptive information, this time referring to recidivism events across the two groups. This information does not translate directly into

the cost models. The process by which this information is translated into costs is described in explicit detail below. Rather, this information is presented in order to help paint a picture of the offending which did occur and to foreshadow the costing process which will be discussed below. In general, it shows the lower number of events the boot camp generated relative to the control group (see Mackenzie and Bierie 2006 for a statistical comparison of recidivism events with these data). As noted above, the cost-benefit is not tied to whether these differences are statistically significant or even favor the boot camp in direction, because the process of using valuation to weight events by the level of 'harm' associated may generate findings that conflict in direction or magnitude with methods which simply test whether 'events' were different.

[Insert Table 14]

Step 1. Program Costs

Program costs were derived from computing (a) the cost per day for treating a single inmate at each facility, multiplied by (b) the average number of days served per inmate at each facility. Subjects at both facilities were assigned a sentence length of 183 days (6 months). However, the total number of "days served" varied because the programs differed in the number and enforcement of rules among detained inmates. Each component of the program costs is described below.

Average Days Served. In all, 13 boot camp inmates were released early, as compared to 11 control program inmates (due to court release orders from judges or

successful appeals). Early release rates were not significantly different across programs. In addition, 19 boot camp inmates were failed from their program, whereas 16 control inmates were failed.[41] Again, the rates were not different than expected by chance alone. Importantly, the control site utilized the "minor sanctions" option of adding 30 day or 60 day sentence increases more often than the boot camp. In contrast, the boot camp was more likely to revoke MAP contracts in full. This meant that despite the roughly equal failure rate, the "average days served" by a boot camp failure was longer than the average control site failures. The "average days served" in each program was a weighted average of the time served by (a) completers and (b) inmates released either early or late. In total, a participant sent to the boot camp served an average of 208 days, whereas the average control inmate served 196 days. The differences were not statistically significant (see Table 15 below). [42]

[Insert Table 15]

Daily Cost per Inmate. An estimate of the day-cost was derived from computing the net annual cost for each facility from budget data reported to the state legislature (2001-2003). The "net" cost takes into account the income production associated with each institution (i.e. sales of good to inmates, contracts for inmate labor with the state,

[41] In addition, 24 control inmates and 3 boot camp inmates were released between 1 and 29 days late. These late releases were not due to disciplinary extensions, because such an extension had to be at least 30 days long. Rather, these represent inmates who were not released due to administrative backups, such as the parole department failing to meet with inmates regarding a home plan, etc. the difference was statistically significant ($p < .05$).

[42] Given these findings, one could reasonably argue that the latter analyses could be constructed by using the same time served (the weighted average across groups). This would be consistent with the conclusion of the significance test. However, the actual time served is used in order to allow that tiny be of noise to be valued because it is in a direction which favors the control site. Thus, the choice of counting a chance occurrence against sets a higher bar for the boot camp to pass in order to demonstrate relative benefit.

telephone surcharges, etc.) which was used to offset operating expenses at each facility.[43] The net total annual institutional budget was then divided by 365.2 days to arrive at the annual "per-day" cost for each year. Next, this figure was divided by the average number of inmates housed at the facility; resulting in the cost per inmate for a single day of housing in each year (on average). After converting to a common metric (2005-dollars), the inmate "per-day" costs were averaged across the three years to generate an overall cost of operating each facility for taxpayers (see equation 1.1 below).

$$\left(\sum_{i=1}^{n} \frac{(Net\_Annual\_Cost)_{ij}}{(365.2 * Daily\_Population_{ij})} \right) / n \qquad [1.1]$$

where,

$i = Year$
$j = Facility$
$n = \#Years\_Covered$

As shown in Table 16, the boot camp cost less per-day than the control site. The differences varied in magnitude by year, but never in direction. Although these may appear to be small values (the difference between $4.00 and $8.00 depending on the year examined), they translate into substantively large values when summed across a full year and a full cohort of inmates.

[Insert Table 16]

[43] Specific information on revenue-budget lines, as well as their use by facilities was obtained through Sue Dooly, a representative of the finance and budget department with the Department of Public Safety and Correctional Services (personal communications, 2005).

Using this information, the final facility cost estimates were computed by multiplying the average day-cost per inmate by the average days served by an inmate sent to each program. See Table17 below.

[Insert Table 17]

Table 17 shows that the boot camp cost *less* per inmate sentence than the comparison site. In fact, it required an average of $470 *less* to send an inmate to the boot camp rather than the comparison site. The savings occurred because of lower expenditures across a number of specific budget lines at the boot camp, including General Administration, Plant Operations, and the greater income generated by the boot camp. Indeed, the per-day cost of operating the boot camp were so much less than the comparison site that the savings were observed *even through* the average sentence was slightly longer at the boot camp. If interpreted at an annual scale, the boot camp may save as much as $329,000 per year relative to the control site, even if recidivism was equal across groups (i.e. assuming a capacity at the boot camp of 700 inmates annually).[44]

This finding is counterintuitive. Most of the boot camp literature, as with most treatment literature, has been rooted in the *assumption* that treatment is more expensive to deliver. However, this was not the case in this data and need not always be the case.

[44] The capacity of 700 inmates annually is a subjective figure, designed to aid interpretation by representing the maximum capacity that would be reasonable for the boot camp. The actual capacity of the boot camp is somewhere around 900 inmates per year. If they chose to deploy all of their resources to MAP-IA inmates (6 month terms) and reserved about 200 beds for staging / evaluation inmates, then the 700 inmate capacity would be reasonable and represents a "maximum" capacity for the effects specified here.

Particularly, the demands of running a large facility may create unique and additional burdens that require more money in administration and operating expenses which offset or overwhelm the savings expected from economies-of-scale.[45] It is important to remember that these are average costs; marginal differences between facilities were not used. In this sense, the present analysis estimates the difference in averages to approximate the marginal difference, but this is not a true marginal cost because some budget lines are fixed.

Earnings lost while detained

Cohen (1998) has argued that, "a prisoner is not generally a productive member of society while incarcerated. The loss in productivity is best proxied by loss in earnings of the offender" (p.16). Many of the subjects in this study worked before coming to prison and would likely have worked had they not been detained. Cohen is suggesting the lost labor had a market value, and was removed from the community because of detention. The two facilities had equal average wages amongst their subjects, due to randomization. However, the boot camp held inmate slightly longer (on average) than the comparison site, which implied the "lost labor" value was slightly greater at the boot camp. Cohen's suggestion implies the boot camp entailed an extra cost on the community; more labor lost.[46]

[45] For example, the Comparison site did not have a program of work release comparable to the Boot Camp's DOT-contract labor. This may be because to the size, mixed-security population, and location in Baltimore City created hurdles to security and administration which would have made such a program unfeasible. Thus, a key source of revenue was missed by the comparison site.

[46] The value of inmate labor is tied conceptually to the value of labor as a whole. If the community was paying these subjects, then their value to the community is represented by that wage. Thus, I am not arguing that the inmates "lost" money themselves and that I should estimate losses from their perspective (the legitimacy of which is in debate). In contrast, the value I attach is from the community's perspective.

Drawing on the Time I survey, 206 out of the 234 subjects reported sufficient information to derive their annual earnings in the year before they were incarcerated. (Inmates were excluded if they failed to respond to the wage item, did not take the Time 1 survey, or reported information in a metric other than what the survey asked.)[47] Out of the 206 who reported data, 48 (23%) reported being unemployed; making *no* legal income in the year before their current incarceration. The average annual salary among those who were working was $20,465. The average salary across all inmates (employed and unemployed) was $15,668.[48] The average per-day value of the labor lost to the community, therefore, was $43.25 (e.g. $15,688 / 365.2). This meant the boot camp generated an average labor loss of $9,011 (e.g. 208.34 days * $43.25). In contrast, the comparison site only generated a loss of $8,486 per inmate served. When factoring these estimates of lost-labor value into the estimate of "program costs," the conclusions regarding operating costs change. That is, the boot camp cost slightly *more* than the comparison site. In this case, the community as a whole (i.e. tax dollars the state would have received as well as the abstract value the community would have attached to their labor) spends $55 more when inmates are sent to serve their MAP at the boot camp, rather than the control site. Again, this refers to the actual expenditures by the state as well as the more abstract concept of "labor lost" to the community.

Some may be skeptical of the appropriateness of accounting for the value of "abstract labor," and the literature is not conclusive with respect to the utility. Most

[47] In the survey, the research team asked for the weekly average earnings in the previous year. Some inmates reported income from drug dealing, which I exclude (a follow up question on the specific job title helped us identify these cases). Others reported their hourly wage (i.e. $6.25 / hour) but I exclude these because I do not know the average number of hours worked.

[48] Cohen (1998) found that the average legitimate earnings of convicted felons hovered at $9,177 per year (2005-dollars), a figure averaged across all inmates in the U.S. The current study's sample was slightly lower-risk than the average inmate sample, which resulted in a higher average labor value.

important, skeptics may argue the value is unfair because the method means a single type of an inmate's community impact is included, while neglecting other reasonable influences (i.e. crimes averted while detained, general deterrence, and reduced risk of danger to community in terms of accidents, poor parenting, or general disorder). The lost-labor value is biased towards making all removal from the community look bad. There is no attempt to capture benefits to a community associated with removal of an offender. This presents an unbalanced tool; using it presents a biased view of community impact. Because of this potential debate, the analyses below were analyzed both ways (with and without incorporating lost-labor value). No substantive differences in findings were found. See Table 18 for program costs which have been corrected for inmate lost-labor.

[Insert Table 18]

It is important to interpret these values in light of the annual "operating impact" they represent. Although these differences may appear small in value, the magnitude of their impact again translates into reasonably large values when interpreted in light of program capacity of the facilities and annual expenses. For example, if the boot camp can process around 700 inmates per year through their facility, the estimates in Table 18 imply the boot camp costs $39,200 per year relative to the control site.

Step 2. Post Release Costs of Recidivism

The third stage of the cost-benefit analysis involves the estimation of post-release behavior. As noted in the previous chapter, this is often considered the most difficult step

in the process. First, it is difficult to include all behaviors that may be theoretically interesting (because of lack of data, lack of resources, or lack of defensible methods for converting those behaviors into monetary metrics). Second, there is some debate as to which benefits are appropriate to count. Some would prefer to limit the analysis only to criminal justice expenses; arguing this is the most conservative and "bottom line" approach to program evaluation (see Aos et al. 2001). Others prefer a full accounting of tangible and intangible costs of crime to victims (see Cohen 2005), and broader social impact of programs (see Nagin 2001). Rather than taking a side on this debate, this section provides a staggered analysis so that increasingly complex models can be compared and policymakers can thus examine results obtained under conditions emphasized by each camp in the debate.

Recidivism Events: What offenses occurred?

Arrests. Estimating the difference in crime-related costs hinges on an accurate and thorough portrayal of the type and amount of offending that was processed by the criminal justice system. The analyses above showed there was a significant difference in the total number of charges generated by the two groups (boot camp and control) over the course of this study. This section delves deeper into this re-arrest information in order to better inform cost differences.

A total of 905 charges were generated by the subjects; 544 from the control site and 361 from the boot camp. Controlling for differences in sample size, this meant there was an average of 3.44 charges generated by a boot camp participant, whereas the control

site generated an average of 4.50 charges per subject served.[49] In Table 19, two trends emerged. First, the largest reductions in charges came from differences in drug offenses. Importantly, the largest differences here were in the realm of 'hard drug' use and sales. Whereas the boot camp subjects were only slightly less likely to generate a "Possession of Marijuana" charge, they generated fewer than half the hard-drug possession charges as controls. Likewise, they generated less than half the number of drug sales charges (i.e. a difference of more than 200%). Second, there were a handful of charge-types in which the boot camp inmates generated a greater amount of the crime. However, the differences tended to be small. The largest was theft, in which the boot camp inmates generated approximately 30% more charges than controls.

[Insert Table 19]

Table 19 reflects all charges, and is therefore a useful tool for understanding the totality of offending across groups. However, this information is most useful when viewed along side the same data after correcting for the impact of 'multiple charging.' Table 20 presents the same data after restricting consideration to the controlling offense (i.e. the most serious charge in each arrest event). This approach is particularly useful because it controls for the possibility that a handful of outlier recidivists in a facility are driving the comparison of charges (i.e. a single person with 10 robbery charges, or a single person with 20 charges of drug sales, etc.).

[49] The same patterns are found if restricting the analysis to the number of charges per recidivist, rather than per inmate served. Here, the average recidivist from the boot camp generated 5.73 charges per arrest, and the average control recidivist generated 6.72 charges per arrest.

In Table 20, the data illustrated the same patterns as above. First, the boot camp generated a total of 114 arrest cases, and the control site a total of 186. These counts were higher than the total population of subjects in each facility because some subjects had more than one subsequent offense. The boot camp generated an average of 1.09 later arrests for every inmate served. In contrast, the control site generated an average of 1.54 re-arrests for every subject served. Converting these figures to reflect the annual impact of the boot camp (i.e. scaling up to ~700 inmates served a year), the difference of 0.45 arrests would translate into 315 fewer arrests if the MAP program was operated at the boot camp rather than the control site. Finally, the data again showed there was a difference in the quality of what offenses occurred (in addition to the difference in frequency). The boot camp generated fewer arrests per inmate than the controls across all categories of offending (violent, drug, property, and other). Although there were some lines in which the boot camp inmates had more offenders (i.e. the boot camp had a single arson case, the control site had none), the major pattern was one which favors the boot camp. Importantly, this was especially true in the case of violent offending.

[Insert Table 20]

Convictions. The same patterns emerged with respect to convictions (data not shown). The subjects generated a total of 102 convictions. Of these, 63 cases were from the control group and 39 were from the boot camp subjects. Or, in other words, the control inmates generated 0.52 convictions for every inmate served, and the boot camp inmates generated 0.37 convictions for every inmate served. When brought up to scale

(annual impact) the difference of 0.15 fewer convictions per inmate served implies that 105 fewer convictions per year to be expected when operating the MAP program at the boot camp rather than the control site (assuming the 700 inmate capacity per year).

Incarceration. The story became more complex when examining reincarceration, with some comparisons favoring the boot camp and some favoring the comparison site. Each recidivist received two sentences. The first was their "maximum" sentence. The second is their "minimal" sentence, which was the maximum sentence minus any suspended time. If the re-incarcerated subjects behaved well, the minimal sentence would be the total time served. If the re-incarcerated subjects misbehaved (either in prison or on parole) then the remainder of their suspended time would likely be applied to the subject. The important point, then, is that neither scenario is likely to be true in *all* cases. Some inmates will fail and serve their maximum time or more (if a new sentence is tacked on for additional offending); some inmates will only serve their minimal time. The best approach for interpreting these data, then, is to estimate the impact of both scenarios in order to paint lower-bound and upper-bound estimates of the cost of incarceration; assuming that truth lies somewhere in between. However, for reasons explained below, the 'maximum time' is likely far more accurate a proxy than minimal time.[50]

[50] As noted in more detail below, emerging empirical research on parolees suggests that the "maximum" scenario will be closes to the truth. This is because parolees who have returned to incarceration at least once before (referred to as "churners") have a substantial probability of failing on the subsequent period of release as well. The national average in this sub-sample is a 75% reincarceration rate, so I expect that 75% of the sentences handed down among the recidivists here will result in a serious enough additional bout of recidivism to justify the application of "suspended time" (see Blumstein et al. 2006; Rosenfeld et al. 2006).

Minimal Jail Time. Turning first to the scenario of minimal time served, the data show the boot camp inmates generated a total of 31 reincarceration spells, representing an average of 0.29 re-incarcerations per subject served at the camp. In contrast, the control site generated total of 45 reincarceration spells, which comes to a rate of 0.37 per inmate served at the control site. The control group, then, generated 0.08 more incarceration spells per inmate served. If taken to annual scale, the boot camp would generate 56 fewer subsequent events of re-incarceration for each year in which it operated as the MAP site, relative to the control site (assuming 700 inmates served per year).

Of course, some subjects may be re-incarcerated for more than one spell over the follow up. For example, out of the 31 reincarceration spells at the boot camp, there were only 22 people who were re-incarcerated (a rate of 0.21 people going back to prison for every subject served at the boot camp). Of these 22 subjects, 7 had two incarceration spells, 1 had three spells, and the remaining 14 had a single incarceration spell after release. Likewise, the 45 incarceration episodes from the control group were generated by a total of 32 people (i.e. 8 who had two reincarceration spells, 3 who had three incarceration spells, and the remainder had a single spell). Thus, the control site generated 0.26 *individuals* returning to prison for every inmate served. The difference still favors the boot camp, but the difference is lower in magnitude (0.26 controls versus 0.21 at the boot camp).

Both comparisons are useful, but for slightly different questions. In the case of the first comparison (total number of reincarceration spells) the data would address whether there was a difference in the total number reincarcerations that will occur as a result of program assignment. If the focus is on state resources, such as processing new

inmates, and paying for their time in prison, then this comparison is most interesting. But the second comparison (absolute number of people returning to prison) is also informative for analogous questions. If, for example, we are considering the impact of incarceration on family disruption, then the absolute number of inmates returned would speak to the cost in terms of the number of families expected to be impacted by re-incarceration.[51] As this study focuses on state resources and does not address these other questions (such as family disruption, the analyses below focus on the first comparison (average number of spells per group).

Looking at minimal sentence lengths, the data indicate the boot camp recidivists had a higher average prison terms among the re-incarcerated than controls. The average prison length for boot camp subjects who were re-incarcerated was 1419 days, while the comparison site was substantially less, averaging 965 days. However, when comparing these two figures, it's important to note the highly skewed nature of the data. For example, the median sentence length for the boot camp recidivists was 183 days, and the median for the control recidivists was 365 – the opposite direction as found when using the "mean" to compare. The point being that the distribution was highly impacted by a few outliers.[52]

---

[51] The models are not mutually exclusive with respect to addressing these questions. In fact, it may be interesting to look into the change in costs associated with chronic re-incarceration versus single events; to see whether there are diminishing costs with additional incarcerations or perhaps additional costs with chronicity, etc. Regardless, the point here is that when speaking broadly, the different presentations of the re-incarceration rates have greater and lesser direct interpretability depending on the question we ask. But both are important.

[52] For example, the boot camp had a single case in which a subject was sentenced to 50 years for an attempted homicide. Recalling that the boot camp generated no sexual assault or actual homicide cases, this makes the attempted murder (a shooting) the most serious offense generated from the treatment site. The control site generated several actual homicides and sexual offenses, each of which is relatively more serious than an attempted homicide. However, none of these other offenders (homicide, rape, kidnapping) form the control site received sentences of more than 30 years. Thus, the boot camp mean was driven by the harsh sentence of a single offender which may indeed be a strange case; an outlier in substance as well as empirical content. If this single boot camp sentence is omitted, the boot camp's "average" sentence

Finally, in terms of time-to-serve, the boot camp generated a total of 43,989 days of incarceration across the 105 subjects; an average of 419 per inmate served at the boot camp. In contrast, the comparison site generated a total of 44,384 days of actual prison time, which was only 367 per subject served. The difference meant the boot camp generated 52 *more* days of incarceration per subject served.

Maximum Jail time. Drawing instead on the maximum time, the data show the boot camp generated 35 incarceration spells, which was an average of 0.33 spells for every subject served. The control site generated 50 spells, which converted to an average 0.41 spells per subject served. The difference of .08 spells again favored the boot camp, similar the same ratio observed when comparing "minimal" time above.

In terms of time-to-serve, the boot camp was favored with respect to total incarceration days generated per subject. Again, this change in conclusions from above was derived from utilizing the 'maximum' time to serve, rather than minimal. The boot camp inmates generated a total of 67,968 days of potential incarceration, which came to an average of 647 days per inmate served at the camp. In contrast, the control subjects generated a total of 81,910 days of incarceration, which was an average of 677 days per inmate served. Thus, in terms of maximum sentences, the data show a slight advantage with respect to the boot camp (30 days less incarceration).

---

dropped to a mean of 858 days of incarceration by each subject sent to the camp; a finding which would reverse the direction of the above conclusions. Likewise, comparing the sentence lengths of recidivists after utilizing a log-transform to normalize the data, the boot camp demonstrates slightly longer prison terms, but the difference was not statistically significant (data not shown).

Step 3. Summarizing Benefits and Costs

The three cost benefit models are summarized below. In each case, the total costs of operating facilities are computed the same way, but the models differ as to which of the post-release behaviors are assumed to cost; whether they count criminal justice losses only, or allow tangible victim losses or intangible victim losses as well. Within each model, the sensitivity of findings to key assumptions made during calculation is presented. These include:

1. Whether operating costs include the "labor lost" correction, or only use the observed net cost of operation,
2. Whether costs should include all subjects charged with a crime (observed losses), or restrict consideration to cases in which legal guilt was verified by the courts, and
3. Whether maximum or minimum prison terms were used to calculate cost of reincarceration

In order to maintain a sense of transparency, and to allow readers to gauge the sensitivity of findings to various assumptions within models, estimates are recomputed in a number of ways for each model. Table 21 provides a summary of the data elements which were used to compute comparisons, as well as the sensitivity analyses across each of the three cost-benefit models. This table is a summary of the different cost-elements described up to this point in the dissertation, as well as how each cost changes depending on assumptions made. The actual numbers in each cell refer to the total value for the boot camp or control group which result from applying (a) the valuations and assumptions described across each of types of cost (police, court, corrections, direct victim losses, or intangible victim losses), multiplied by (b) the actual offenses for each person in each

group and then summed within each group. Thus, these 'elements' will be used in the following sections to compute the total cost of recidivism for the two groups according to the three different models described: Criminal Justice, Direct Loss, and Full cost-benefit model.

[Insert Table 21]

Criminal Justice model

This section presents the first and most conservative model for estimating post-program recidivism costs (see comments regarding validity on p. 17 above). It begins by examining the direct losses to police for making an arrest, courts for processing them, and corrections for administering sanctions. This analysis is considered "conservative" because it only allows crimes to be valuated according to the tangible criminal justice expenses associated with administering justice, and the Criminal Justice expenses are far smaller in magnitude than victim losses (on average).

Police. Drawing on the above estimates for the monetary loss associated with making an arrest, values for each recidivism event were computed.[53] Findings were replicated using two methods for 'counting' police losses. The first valued the cost as it occurred (an "observed costs" model). The second restricted computations to only include those crimes which were later verified as deserving arrest (i.e. a finding of 'legal'

[53] Consideration was limited to that associated with the controlling offense. A cost of each charge in an arrest was not used, for example, because this would assume a separate police officers time and resources were used for each offense. Although this is possible in some cases (i.e. multiple robberies by the same offender), it is likely a very rare event. The procedure here results in a conservative estimate of the cost of crime with respect to the boot camp benefit (which averaged fewer charges per arrest event).

guilt). In the latter case, the policing costs were held constant at zero for crimes which resulted in dismissed or dropped charges. As noted above, the value of using both approaches is that it allows one to create a range of potential losses (high and low), within which the true costs lay. However, the true costs were likely much closer to the 'observed' losses than 'legal guilt' only. [54]

Observed costs. Attaching the individual "best" values from the "making an arrest" costs to police found in Table 9 resulted in an estimated loss of $78,864 in 2005 dollars for the 186 arrests generated from the control site. This translated into an average cost of $424 per arrest, and an average of $652 in police expenses drained by the average control inmate (i.e. $69,393 / 121). The average cost per inmate was higher than the average cost per arrest because some inmates were arrested more than once, raising the average loss for the group as a whole.

In contrast, the boot camp's 114 unique arrests drained a total of $39,342 in police spending. Again, this computation drew only on the single most expensive charge as the controlling offense reflecting costs to police. The total translated into an average of $345 per arrest, and an average loss of $374 per inmate served. The difference in cost per arrest

[54] Measuring the observed costs is closer to the truth than restricting ourselves to legal-guilt (although truth likely lies somewhere in between). In short, this is because it is safer to assume that the people listed as "dropped charges" in our data set (i.e. nolle pros, STET, etc.) were indeed guilty, than it is to assume they were not. They could have had charges dropped due to technical problems with paperwork, with evidence collection (the refusal of witnesses to step forward), or because their transgression was handled informally (i.e. community service, treatment, a change in parole standards, etc.) rather than a formal conviction (see Kingsnorth et al. 2002 for an example). In short, the nature of establishing legal-guilt (rather than actual-guilt) is such that a healthy number of avenues exist for having a case dropped other than innocence (Forst 2006). The point being that the observed losses are likely real in these cases. As the state did pay for them, it is more fair to assume this was due to actual differences in subject behavior (i.e. recidivism) than to assume it was not. However, in order to gauge the impact of this assumption, we repeat all analyses after changing our assumption and only including the "legally-guilty" and holding all other CJ costs at "zero" (i.e. all policing and court costs if the case later resulted in a STET, Nolle Pross, etc.).

was found because the boot camp recidivists were committing less serious offenses than the control group recidivists.

Therefore, the boot camp generated $278 less in police losses for every inmate served ($652 – $374). Again, if this figure was translated into an annual operating metric (assuming 700 inmates served per year), the difference implies that the boot camp could save $194,600 for police departments alone each year the MAP program was managed at the boot camp rather than the control site.

Legally verified guilt. Our second approach to calculating police costs is exceptionally conservative, but useful in order to (a) establish the sensitivity of findings to extreme assumptions and (b) paint a lowest-bar figure when interpreting results. When restricting consideration of police losses to only include those arrest events that later result in a finding of legal guilt, the data showed the boot camp generated a total of $13,672 in losses, which was an average of $130 per subject served. In contrast, the control site generated a total of $26,488 in police losses, which was an average of $219 per inmate served. The difference ($89) was not as large as found when relying on observed losses, but the direction was the same (favoring the boot camp). Importantly, the loss still translates into a substantively large sum if interpreted in the context of annual savings over a full facility (approximately $62,300 per year).

Courts. Out of the total pool of arrested-cases, only a portion resulted in a conviction. In the case of a dismissal, a single arraignment was assumed to have occurred. In the case of dropped charges (i.e. nolle pros or STET), two hearings were

assumed to have occurred (arraignment and a major hearing). For "not guilty" findings, three hearings were assumed (arraignment, minor hearing, and major hearing). In the case of a conviction, four hearings were assumed (arraignment, minor hearing, major hearing, and sentencing hearing). In each case, these are usually the minimal number and type of court hearings to generate the court outcome in question. Thus, the cost to the courts for each offense is a function of outcome, rather differentiated across crime types.[55]

Once again, two analyses were conducted. First actual 'observed' losses endured by the system was examined. Second, consideration was restricted to losses under the stringent assumption that all costs are zero unless the courts later determined legal guilt for the arrest.

Observed Costs. Out of the 186 arrests in the control sample, 27 (15%) were dismissed, 90 (48%) were dropped, 6 (3%) resulted in a finding of "not guilty," and 63 (34%) resulted in a conviction. Using the above estimates of court losses, this meant the control site generated a total of $380,994 in court expenditures; an average of $3,149 per inmate served.

The boot camp's 114 unique arrests resulted in 40 (35.1%) subjects found Guilty, 4 (3.5%) found Not Guilty, 59 Nolle Prossed or STET (52%), and 11 (9.6%) Dismissed. This generated a total of $240,686 in court losses overall, which was an average of $2,292 in expenses for every inmate served (see Table 22).

[Insert Table 22]

[55] Unfortunately, the state of MD was not able to provide data on detention during trial, public defender use, or information on the number of hearings or whether guilt was obtained through jury trial, bench trial, or a plea agreement.

The data showed that the state of MD saved $857 in court expenses for every inmate who served their MAP at the boot camp rather than the control site ($2,292 - $3,149). If taken to scale, the data suggest around $599,900 would be saved per year if the boot camp rather than the control site served a capacity of 700 MAP contracts annually.

Legally verified guilt. Once again, computations were replicated while limiting consideration to legally verified guilt alone. In this case, there were 40 cases of guilt at the boot camp for a total loss of $150,520 ($1,434 per inmate served lost due to court expenses). The control group generated $237,069 for their 63 guilty verdicts, which came to $1,959 in court expenses for every subject served at the control site. Thus, the difference under the 'legal guilt only' assumption still amounted to $525 (i.e. $1,959 – $1,434), in which the boot camp was again favored. Brought to scale, this reflects a potential savings of $367,500 annually for the Maryland courts if the MAP program continued to operate out of the boot camp rather than the control site.

Sanction administration. Finally, as noted above, there was a difference with respect to reincarceration use between groups, and the difference changed depending on whether analyses examined minimal or maximum terms. In this section, the issue is revisited by comparing the estimated costs of incarceration for the two groups under both scenarios. Although there is no ex-ante method of knowing which inmates will serve maximum or minimum time, the national trend suggests that 75% of repeat parolees

(such as the re-incarcerated in this study) will be rearrested upon release from a second incarceration.[56] Likewise, each additional arrest thereafter can be expected to generate a 75% additional fail rate (Rosenfeld et al. 2006). In as much as being repeatedly arrested is grounds for parole revocation (i.e. serving remaining suspended time) or the potential of additional sanctions, then the 'maximum' term listed here should be closer to the actual costs that will occur.

The analysis first focused on the "minimal" time ordered (the maximum sentence, minus the suspended time). In estimating the day-cost of sentences, the sentences were divided into four groups (Jail, Minimum, Medium, and Maximum security). As specified in the methods section above, a day-cost was attached to each sentence based on the length and nature of the offense by approximating Maryland DOC classification protocols.[57]

Focusing on the "minimal" time to serve, the data showed the boot camp inmates generated a total of $2,520,567 in incarceration costs for the 105 subjects. The average cost of the boot camp, then, was $24,005 of later incarceration for every inmate served (assuming all served their minimal time only). In contrast, the control site generated $2,694,627 in prison costs across the 121 subjects, which amounts to $22,270 of average incarceration costs for each subject served at the controls site (minimal sentence time

[56] That is, the national average derived from Langan and Levin (2002) report that around 60% of all parolees are rearrested within three years. However, the recent work of Rosenfeld and colleagues (2006) demonstrate that this "average" is composed of two groups. First time inmates recidivate at a much lower rate upon release (~50%). Inmates released for a second term recidivate at a rate of 75% within three years. The rate stays constant at approximately 75% for each additional release. Averaging these two values (first and repeat incarcerations) generates the familiar recidivism rate of 60%.

[57] To be clear, the original security score for each of these inmates was low and contained no violent convictions (which was how they qualified for the boot camp in the first place). Combining this information regarding criminal history, along with current offense information, allows us to make reasonable approximations with respect to security level for the present convictions.

assumed). Thus, the boot camp cost an average of $1,735 *more* with respect to latter incarceration for each subject served (assuming minimal sentence lengths).

In terms of "maximum" sentence terms, the boot camp generated a total of $3,906,457 in prison expenses, amounting to $37,204 per inmate served. In terms of maximum sentences, the control site averaged $4,861,087 of prison costs. This would amount to $40,174 per inmate served at the control site. The difference favors the boot camp ($2,970 saved). As noted above, there is strong reason to believe that the latter figure is a better approximation of the expected costs, because most inmates (75%) will serve most or all of their maximal term. In as much as these 'repeat failures' also generate additional convictions when released early, then the maximum time may be served as well as new and additional prison time. In this case, the benefits of the boot camp would be greater than the $2,970 in this analysis (all else constant). For these two reasons, the latter figure is likely the more accurate.

The Criminal Justice model is the most restrictive; representing a conservative evaluation of the boot camp because it generates the lowest monetary value of the observed difference in recidivism. That is, losses observed by the criminal justice system are included, and all other post-release costs (i.e. tangible victim losses, intangible victim suffering, and non-crime benefits) are all held constant at "zero" value. See Table 23 for a summary of cost benefit models, as well as illustration of the sensitivity of results to the 'legal guilt' assumption.

[Insert Table 23]

In terms of computational assumptions, the most accurate 'Criminal Justice model' derived from including (1) the observed operating costs of facilities, (2) the observed criminal justice losses, and (3) the maximum potential prison terms (for reasons discussed above). In this case, the total cost of sending one subject to the boot camp was $52,298. This included the cost of operating the program, as well as the cost of post-release behavior which generated losses to the criminal justice system. Under the same conditions, it cost $56,998 per inmate served at the control site. Again, this is an average total cost reflecting the "statistical" inmate. Thus, there was an average of $4,697 less cost for every inmate sent to the boot camp rather than the control; a savings of 8% per inmate.[58] The net social value of $4,697 would generate a difference of approximately $3.27 million in constant-2005 dollars of savings over the control site per year (assuming 700 inmates served annually).

As discussed earlier, a second way to summarize the difference is to examine the marginal cost-benefit ratio; identifying how much return is received for each additional dollar invested in the boot camp rather than the control site. This data set presented a unique scenario in the cost benefit literature, because computing a cost benefit ratio required the estimation of the benefit derived from spending less – a scenario that, to my knowledge, has not been encountered in a formal evaluation. The uniqueness of the scenario, in fact, made the computation of a cost-benefit ratio impossible from the perspective of the boot camp – because there was no marginal cost.[59] Not only would the

[58] If I alter the assumptions slightly, by restricting our analysis to the legally-guilty recidivists, boot camp benefit decrease; indicating the state saves just over $755 (1% difference in total costs) for every inmate sent to the boot camp. Although it is small difference, it does amount to a large figure when interpreted at scale (i.e. annual serving capacity).

[59] More specifically, recall that the computation involves computing the total marginal benefits of a program, divided by the marginal costs. In this case, the "negative value of cost" could be interpreted as a

summary measure lack substantive meaning, but it could not be mathematically computed because it required division by zero.[60]

Sensitivity. As noted above, there were several assumptions that could be altered which would change the component costs entered into the model. For example, if "labor-lost while detained" was allowed to impact operating costs, the boot camp cost $55 dollars more to operate than the control. If this adjusted program cost figure was used, along with the same post-release costs as calculated above, then the boot camp would generate $4,227 fewer dollars in later criminal justice losses despite the added $55 dollars in expense to run (per subject served). Thus, regardless of the assumptions made with respect to labor-lost values, the difference in post-release costs to the criminal justice system remained.

Under the most plausible assumptions, the findings remain unchanged: the boot camp is favored and the "benefits" of operating the facility are greater than any costs. However, this is not always the case. If the computational assumptions assumed that (1) labor-lost during incarceration matters when computing operating costs, (2) that all recidivists who were re-incarcerated will server their minimal time only and then generate no more costs afterwards, and (3) that only the legally guilty recidivists "count" (i.e. all STET, Nolle Pross, and Dismissed cases are assumed to represent false arrests), then the control site would be favored. Each of these are implausible, and relying on all three of them together creates an unreliable comparison. However, generating this

---

benefit to the program – and thus added to the numerator. However, the "cost" it technically zero after this subtraction, because there is no cost. This would result in division by zero.

[60] This problem is generated because the lower costs are viewed as a 'benefit' and added to the numerator. At the same time, the value is subtracted from the denominator, resulting in a denominator that is the subtraction of two equal values (and must, therefore, be zero).

analysis is useful in order to gauge an extreme boundary on what our expectations could be.

In this case, the boot camp would cost $47,055 in total expenditures per inmate served, and the control site would cost $45,917. The difference would favor the control site, implying $1,138 saved for using the control site. The difference is almost entirely driven by the assumption that recidivists in our sample will only serve their minimal time (i.e. succeed on parole and never be re-incarcerated.

This figure may appear a point of concern, as it represents large consequence if the assumption made in the main analysis are true. However, two qualifiers should come to mind. First, it is unlikely to be observed given the unreasonable nature of the assumptions necessary. Second, the difference driving the finding (the potential $1,138 potential savings for the comparison site if the above assumptions are true) is completely offset in the more thorough models below which account for victim losses resulting from crime. In these latter models, all analyses favor the boot camp regardless of which assumptions are made (operating costs method, prison-time assumptions, or legal guilt). Using the latter models (direct loss model or full cost benefit), the boot camp is favored no matter how creatively and unfairly assumptions are manipulated in efforts to favor the comparison site.

In sum, the "Criminal Justice" model is endorsed by some because of the conservative nature of the estimates. Even when we have evidence that a program is creating benefits relative to the alternative option, we only allow a very small set of factors to be "counted" when weighting that benefit. If a program demonstrates cost-efficiency under this scenario, then we have strong evidence that the program will save

money because the assumptions are extremely biased toward favoring the control site. (Thus, a program that overcomes this bias is thought to be extremely well defended as a valuable policy option). However, a major draw back to the "Criminal Justice" model is that its conservative nature means that the estimates of crime costs are unreasonably small; they gain relative reliability over other methods at the cost of relative validity. The method ignores very clear and defensible costs enacted by recidivists, such as direct losses endured by victims. The literature has generated strong, redundantly validated tools for gauging the tangible losses of victims from crime. Including these losses creates a more accurate portrayal of program costs; as described below in the "Direct Loss" model.

The Direct Loss model

In this second model, tangible victim costs are added in with the criminal justice losses from above; those costs which are average losses endured by victims of crime across various offenses. Again, these are direct, tangible, and easily measured losses in which there is little controversy regarding validity of values. For example, it includes the average amount of money stolen, injuries received, work missed, etc. across each crime type (see Miller et al. 1996).

Drawing on the controlling offenses, the data indicated the boot camp subjects generated a total of $218,832 in direct victim losses across their 114 crime events. This came to $2,084 in tangible victim losses for every inmate served at the boot camp. In contrast, the control group generated a total of $4,046,091 in tangible losses to victims, which is $33,439 in victim losses for every subject served at the control site. The data

show a savings of around $31,355 in direct victim losses for every subject served at the boot camp rather than the control site. The difference between facilities is large; driven by the fact that homicide and rape cases generate monetarily expensive amounts of injury. The boot camp generated none of these crimes while the control subjects generated several. Rather than being the exception, this was the rule: the comparison recidivists were more harmful than the boot camp site.[61]

Following the approach used in the Criminal Justice model, the analysis was repeated after limiting consideration to the "legally guilty" only. In this case, the data indicated the boot camp generated a total of $32,592 in victim losses, which averaged $310 per inmate served. In contrast, the comparison site generated $1,329,662 in losses; an average of $10,989 per inmate served at the control site. The expected savings in this restricted model is $10,679. The difference is smaller than the "observed" cost computation, but still a substantively large difference (particularly when interpreted in the scale of "annual capacity"). Again, this model assumes that any case in which a crime generated a dismissed, stet, nolle pros, or not guilty court outcome was due to actual-innocence and the subject in question was not in fact associated with the crime.

Drawing on the 'best' assumptions, as defined above, the data indicate the boot camp costs $54,382 per subject served. This includes observed cost of running the program, as well as the observed losses to the victims and the justice system (no legal-guilt restriction). It also draws on the assumption that the maximum prison terms assigned to the re-incarcerated are the best proxy of likely expenses. Under these same

[61] Again, this is indicated in two ways. First, the rate of recidivism for the more (substantively) harmful offenses were greater at the comparison site than the boot camp; offenses such as kidnapping, rape, murder, etc. In addition, the fact that officers and the courts attached more charges per arrest on comparison site recidivists than boot camp recidivists is a lose indicator of the harshness of the criminal act. That is, the more "crime" or "harm" done, the more charges we would expect to be levied.

assumptions, the comparison site cost $90,664 per subject served. The difference ($36,282 per subject served at boot camp rather than the control site) is substantial and remains so regardless of the assumptions made.

For example, if all possible assumptions were made which would manipulate the data to favor the control site (i.e. labor-loss included in program costs, consideration restricted to legal-guilt cases only, and minimal jail time assumed), then the boot camp would cost $47,365, the control site $56,906, and the difference would still imply a savings to the community (state and citizens) of $9,541 for every inmate sent to the boot camp rather than the control site.

Full Cost Benefit model

As noted in the literature review above, there have been several models used to estimate costs and benefits within criminological program evaluation. The gold standard is the 'full' cost benefit analysis. It is the most controversial, because the methods used to derive monetary values for abstract events like "mental anguish" or "suffering" are technically difficult to follow, and may be more fragile in terms of assumptions and generalizability to the non-average situation. Regardless, the technical savvy and cross validated estimates in the economic literature have overcome most criticisms; resulting in far more skepticism hurled at studies which omit intangibles than those that include them.

Focusing again on the controlling offense for each arrest event, the data indicate a substantial difference in the intangible costs of crime generated by the two facilities. The boot camps 105 subjects (and their 114 re-arrests) generated a total of $706,180 in intangibles such as average pain and suffering associated with each crime type. This

came to an average of $6,726 in losses for each inmate served. The control facility, however, generated a total of $7,910,208 for intangible losses, which was $65,374 per inmate served. The difference is enormous, representing a savings of $58,648 in terms of pain, suffering, and other intangible costs of crime for serving the MAP contract at the boot camp facility rather than the control site.

Drawing on the 'best' assumptions, as detailed above, the Full Cost Benefit model showed the boot camp to cost $61,108 per subject served, while the control site ran $155,808 per subject served. The difference is again large, implying that $94,700 was saved for each subject sent to serve their MAP contract at the boot camp rather than the comparison site.

Once again, the estimates were recomputed under after restricting costs to events to those in which legal guilt was later established. Under these restrictions, the boot averaged $52,532, and the control site averaged $88,154. The difference is still large and favors the boot camp ($35,622 saved). Likewise, the benefit to the boot camp is large even if we alter all possible assumptions in order to artificially favor the comparison site (as described above). Even under these unlikely circumstances, the boot camp would costs an average of $30,392 less than the comparison site ($48,344 total costs for the boot camp, versus $78,736 total cost for the comparison site).

CHAPTER 4: DISCUSSION

This study focused on summarizing and comparing the costs and benefits of a specific policy choice facing the state of Maryland: To operate the MAP program at the boot camp, or at the control site. In pursuing this end, the analysis has drawn upon a randomized experiment comparing the two policy options and calculated the cost of operation as well as post-release offending generated by the competing policy options. Because of the strong research design, the observed differences in behavior and costs could be attributed to program with more certainty than observational or quasi-experimental designs. From this comparison, several interesting findings emerged.

First, the data suggested the boot camp generated fewer charges per arrest and fewer arrests over all, at least in terms of directions of mean differences. The process of cross-referencing multiple databases to generate recidivism data generated greater reliability in the recidivism measure than would be observed if relying on police data alone. This paper does not include statistical analyses of these differences (see Mackenzie and Bierie 2006 for statistical comparisons). Rather, these data were simply a starting point for cost-benefit weighting processes.

Second, the analysis compared the cost of operating each policy option. Rather than relying on the state budget computation of 'day costs' for each facility. A 'net day cost' figure was generated by accounting for income, as well as expenses, endured by each facility before computing a rolling average of costs. This computational decision was important because it more accurately captured the actual costs of facilities from the perspective of the department of Public Safety. Likewise, it gave credit to facilities for

offsetting their own expenses. The data indicated the boot camp actually costs less to operate than the comparison site. This was a novel finding in the program evaluation literature in general, and especially in terms of criminal justice evaluation. Not only did it impact the present evaluation, but the finding may serve to encourage policymakers to question basic assumptions regarding policy costs in the future. For example, policy makers (including those this dissertation was designed to serve) operated under the assumption that a the boot camp must cost more than a traditional prison; assuming economies of scale and reduced services *must* translate into reduced costs. This was not the case in these data. Although beyond the scope of this analysis, it seems plausible that the smaller facility operated more efficiently and with more safety and control which allowed them to pursue contracts for inmate labor which more than offset a higher staff to inmate ratio. Likewise, the more efficient management may have used fewer resources in terms of staff overtime and sick leave. Future research should delve more deeply into this finding, as the policy implications would be paramount for budgetary reasons alone.

This finding of lower operating costs was sensitive, however, to the assumption that 'community lost labor' was irrelevant. Once program costs were recomputed after accounting for this abstract cost to the community, the data indicated the boot camp did indeed cost more ($55 per subject served). This higher cost was derived from the slight increase in time served at the boot camp. This sensitivity should be interpreted cautiously, because the addition of 'labor-lost' is controversial. It does not serve as an unbiased proxy of 'community impact' of incarceration. Rather, it only reflects the 'good' these subjects were doing in the community. Thus, removing them longer is assumed to be a 'loss' for the community. Although it is fair to argue that offenders may

provide benefits to communities, it is unfair to ignore the possibility that they also had negative impacts (i.e. offending, or risky behavior). In as much as subjects' impacts on the community were truly complex (involving both positive and negative impact), the use of a single proxy for that impact which only taps the 'good' impact will necessarily bias the findings. For this reason, I lean towards the 'observed' costs of operating the two facilities as the more accurate figure.

Third, the data indicated that the boot camp cost less than the control group in terms post release offending. The boot camp was favored regardless of how computations were made within the 'direct' and 'full' cost benefit models; no matter how extreme or unreasonably the deck was stacked against the camp. The computations based on the most reasonable assumptions within each of these two frameworks demonstrated that Maryland citizens gained tremendous savings from the utilization of the boot camp, rather than the control site for the MAP program.

The 'criminal justice' model also supported the boot camp, but the savings were sensitive to extreme assumptions. In particular, if the computations assumed that all subjects would serve the minimal sentence ordered by the court only, the comparison site was favored. However, the plausibility of this finding is tenuous because the 'time served' assumption is unlikely to be true. Likewise, the model in general is less accurate because it required the assumption that victim costs were zero (see Cohen 1988, 2005).

Finally, it is important to note that the approach used here (estimating multiple models) had a secondary benefit, a benefit which may appease readers skeptical of the valuation literature. These skeptics would likely want a confidence interval rather than a point estimate placed around valuations of crime costs, particularly with respect to victim

losses (intangible and tangible). For example, they may suggest "the cost of rape can't possibly be nearly $100,000 in losses, no matter how much we detest it." In this dissertations defense, in fact, this statement was presented as a critical flaw in the cost benefit literature - the unreasonably high costs of these full estimates. Although the empirical literature (and many rape victims) would disagree with this criticism, it is grounded in a reasonable statistical question: how precise is this estimate and what is the 95% confidence interval for each type of crime? In short, a cynic would prefer a statement such as "a rape generates between $35,000 and $200,000 in losses, and here are the computations repeated if (a) I assume the value of the lower bound is true, and (b) here they are again if the higher bound was true."

To date, the valuation literature has not provided confidence intervals around point estimates that are presented as 'plug in values" (see Cohen 1988, Miller et al. 1996, Cohen 1998, Cohen 2005, and French et al. 2000 for examples of studies which do not use them because they are not available). Importantly, they are not available because the key valuation literature has not been derived from a 'typical' sampling/estimation process from which confidence interval computations would generate meaningful estimates to use. More specifically, the key valuation literature derives from including estimates of average values from multiple sources and methods, and the underlying distribution is not known. For any particular crime, an organic process occurs in which the component parts of the loss (i.e. lost wages, hospital fees, wage changes, pain and suffering) are derived from multiple sources and sometimes from multiple methods. That is, multiple sources may be used to generate an estimate of each component, and then different sources or methods to derive another component, etc. The rape-valuation derived by

Miller et al. (1996) involved the use of jury award data for pain and suffering, public health data on hospital fees, lost wage and average injury information from victimization surveys, etc. Authors who can generate estimates of average costs derived from average injuries have found that generation of confidence intervals from multiple-source multiple-methods data are not meaningful, because the usual methods of confidence intervals do not necessarily generate a meaningful boundary that reflects the accuracy and consistency of the point estimate (the point of confidence intervals) rather than being partially confounded by differences in methodology or sources. Thus, I omit them because there are not meaningful values to use to accomplish this end in the published literature.

This is not true in all cases, of course. When a valuation study uses a sole method to illustrate one component cost (i.e. jury based estimates of a rape victim's psychological-suffering), then we have the potential to derive a confidence interval for that estimate. Although these smaller and specific studies of valuation feed into the generation of "plug in" values as found in Miller et al. (1996) or Cohen (1998), the papers providing holistic plug in values tend to provide point estimates only for the reasons stated above.

That being said, the dissertation provided an analogous sensitivity test, even more stringent than placing 95% confidence intervals around estimates. Again, it was a concern for these very issues that lead me to think critically about the meaning of these point estimates and how to best provide sensitivity analyses for readers (an analogue of the confidence interval). The best approach, I decided, was to invoke a labor-intensive process by which each analysis was repeated under different 'costing' assumptions; by re-computing all analyses after making assumptions about the values of crimes that were

substantively and theoretically guided, but analogous to a confidence interval to generate lower bounds.

Each of the three "models" used to repeat the analyses was, in this sense, a sensitivity analysis of lower bound figures. In the CJ model I was really asking, "What would we find if I assumed all victim costs were zero?" In the second, I asked, "What would we find if we only included exceptionally small values for victims?" In the final estimate, I asked, "What would we find if I used the full and most accurate point estimate of victimization impacts?" Note, for example, that in the CJ model, I placed the victim-value of Rape at "$0.00," in the Direct model it was, "$6,278," and in the Full it was "$107,000." In short, the present analysis was generating a more extreme sensitivity analysis than we would likely have encountered if we had placed a 95% confidence interval around the $107,000 figure. The information derived from a confidence interval-based sensitivity analysis would be redundant with the findings in this paper which used larger bounds than a CI would likely have generated. Thus, even if confidence intervals were available, such an analysis would also show that the effects are consistent and favor the boot camp facility across models.

Again, this process generated the lower bound estimates of costs that are analogous to a confidence interval around valuation values for each crime type, although my approach was an even more stringent test because my assumptions were likely generating costs beyond a simple 95% interval estimates lower bounds, whatever they would be. I chose to concentrate on providing lower bound estimates because I doubted that cynics were going to suggest the cost of Rape was too low, I knew the criticism would fall from people who were unsure crime was quite as costly as econometric methods suggested.

Indeed, this hunch bore out in the defense process. However, an additional reason to concentrate on lower-bounds was because my theme has been to set the deck against the boot camp when faced with subjective coding decisions. In this case, the boot camp had no rapes at all, and the control site had several. To amplify the estimated costs (even if true) would only be redundant with all the other analyses in this paper in showing yet again that the boot camp was favored.

A related question has to do with sensitivity of findings to extreme types of findings. For example, we may ask, "Are these findings being driven by the inclusion of a handful of expensive 'homicides' found in the control group, while the treatment group had none?" Thus, the question would be whether we should exclude them from consideration when generating estimates. This question is worth examining explicitly.

The premise of excluding these cases from consideration would derive from the assumption that there is something unique or strange about the homicide crimes, that mean the subject would have killed regardless of the facility assignment, and that these strange cases were all assigned to the control group by chance alone. Thus, they would have killed anyway and we should not assign a relationship between these acts and facility assignment.

This is a troubling stance for several reasons. First it presents an assumption that the crime of homicide is somehow unique from other offending, outside the purview of treatment interventions, and essentially deriving from 'sick' individuals. This is view is hard to defend from the literature, empirically or theoretically (especially if this is seen as the general case). I don't think this is a fair view of homicide as a crime, which may not be that unique from other offending (at least in many homicide events, although there

must certainly be diversity). After all, if the premise is that they were innate killers, then why haven't hey killed earlier? Why wait until age 25-30? Rather, its important to remember that these subjects were primarily drug dealers from Baltimore City, and those in the control group who continued or accelerated their involvement are likely imbedded in life styles that involve black market violence to regulate their industry. In short, I expected homicides in these data among those who continued their 'lives in crime' because there is a real risk of violence within this context and some of those acts will purposely or accidentally lead to death. In fact, I would be troubled if there were no homicides in the control group – rather than being troubled because there are. I am more comfortable arguing that we would have expected a handful of the boot camp subjects to have committed homicides had they not been sent to the camp for these reasons than assuming there were a couple of subjects that were inherently murderous and would have been off killing no matter what we did with them.

Likewise, I am uncomfortable with the methodological and ethical implications of post-hoc alterations of data and randomization; with selectively rejecting the assumptions on which randomization is based. Why not remove all the rapes too? How about robbery? We could tweak these data endlessly until the findings were altered, if we desired. But that requires assumptions that are harder and harder to justify, and leave us with a comparison of groups which is hard to interpret, and a question as to the role of post hoc manipulation in social sciences.

Regardless, I re-ran all analysis with these offenders excluded and the direction of results did not chance. The magnitude decreased, but the differences between facilities were still substantively large. This should be expected from examining Table 19 in

particular. The story in these data were consistent across offenses; the boot camp was almost universally favored in terms of offending per individual. And the differences were even more pronounced across the more serious offenses. In short, the consistent story in these data was one of a marginal difference in rates of offending – not a "similar picture" accept for a few outliers that happened to carry high values. We should not expect this exclusion to matter much, and it did not.

Limitations

The present study was not without limitations. First, some cost data was not available. For example, there was no information on the use of Public Defenders, intermediate detention (detention during trial), or the number of actual court hearings. In the later case (court hearings) I was comfortable assuming the costs varied according to court outcomes rather than using the missing information on 'count of hearings,' because the actual count of hearings is often a function of processes *other* than the seriousness of the recidivism event (i.e. legal complicatedness, a defendant who believes he is innocent, etc.). However, it would be preferable to have these data and test the sensitivity of the findings to other coding and computational schemes rather than assuming the method used was the best. However, it is likely that each of these three lines of data would also favor the boot camp facility, because the rate of offending and the seriousness of recidivism were lower at the boot camp.

Likewise, there was no budgetary data on ISP parole supervision. The boot camp group was monitored on ISP more often than controls. ISP likely cost more because monitoring costs money, and ISP has more of it (more visits, lower case loads for

officers, etc.). In addition, it likely generates more technical violations which may cost more because (a) a technical violation hearing is held, and (b) sanctions may be applied which may also cost money. However, there was no budgetary data available that allowed me to estimate the cost of parole (either general monitoring or cost of violation hearings). More importantly, it was unclear whether this would be appropriate to count as a boot camp cost, as the goal of this paper was to estimate the cost of the boot camp prison versus the control prison, and post release crime. The parole monitoring was a separate intervention being applied, but was not part of the boot camp (indeed, the state does not want this confound; in their eyes the facility versus parole costs are distinct and separate programs of interest). ISP may or may not be used with future boot camp subjects (as with other types of parolees), but appears to be a separate question when we are referencing the costs and benefits of sending inmates to serve their time at the boot camp or the control site in the future.

In addition, the state did not provide good data on what happened as a result of technical warrants, other than (1) counts and (2) the designation of outcomes as either "nothing – continued on parole" or "parole closed" as outcome possibilities. There were not many technical violations filed, although the boot camp had more 'per-subject.' Regardless, most resulted in outcomes were "nothing – continued on parole," and thus few additional costs. Parole officials told me that even in the case of the rare "closure-outcome" or "revoked" parole, the subject actually does not receive a sanction of any sort (such as jail return). Instead, their parole case is closed with the label of "closed unsuccessfully" and they are simply set the subject free. The 'sanction' is that the record of having been noncompliant remains in subjects file. If they end up recidivating in the

future, then this past noncompliance can be used against them during a future sentencing hearing.

A key reason to chose this option rather than a 'return to jail' appeared to be tied to the fact that most of these subjects only had a year or so of parole to start with, and by the time they 'violated' they only had a few months or weeks left on parole. The state doesn't generally want to spend money on incarceration for a few weeks/months because of technical violations because the processing fees are enormous and the expected deterrence value is small. Instead, they simply close the case and tell the person there is a record which will be used against them if they ever return to court.

If these omitted costs were included in the calculations, the benefits of the boot camp would diminish, although the impact would likely be small. The boot camp had 50 subjects on ISP parole, while the control site had 37. Although there were more technical violations at the boot camp group, there were only a handful of these that occurred, and only a handful that resulted in any 'revocation' status. In addition, there is likely only a handful again of these revocations which engendered costs in the form of sanctions, these sanctions were likely only for short durations until the parole term ended, and so the likely impact would be small.

In contrast, other omitted variables would likely have favored the boot camp. In other research examining these data, we identified additional benefits to the boot camp, such as a near 400% increase in GED achievement for boot camp subjects relative to control subjects. Cohen (1998, 2005) draws on educational literatures and notes that the direct benefits in terms of lifetime earnings differentials between dropouts and those with GED's is roughly $300,000, with an additional $90,000 difference in fringe. The

estimate is even higher if the figures include quality of life changes (the 'intangibles from the education field). This dissertation omits these benefits as well, although including these differences would likely dwarf the difference in ISP costs, favoring the boot camp.

Second, the paper was limited in the specificity of data obtained on actual crimes. For example, it would be helpful to have data on the type and amount of drugs involved in drug-related recidivism. It would be plausible to extract valuations of drug offenses if these two pieces of data were available (see MacCoun and Reuter 2002). However, the court, police, and parole data did not contain this specificity. More specificity may have been useful, as it seemed likely that the trends favoring the boot camp in these data would have also played out in terms of size and types of drug involvement among recidivists. For example, recall from the analysis of charges above (Table 19) that the control site had nearly twice the proportion of hard drug sales among their subjects relative to boot campers (both in terms of hard drug possession and sales).

The project would have benefited from the observation of more detail on other crimes (i.e. actual police reports) for a second reason. The cost-benefit literature often relies on official records, such as police raps sheets indicating an arrest occurred, or court records indicating the outcome of arrests. However, these data are often obtuse with respect to how many victims were involved in a crime, or how many victimizations of a single victim occurred. This data can be critically important in terms of victim costs. However, the literature to data has been silent on the topic of charging and multiple-charge valuation. By examining police reports rather than court data, value could at least be attached to each victim in an event. If the future valuation literature takes this topic to heart, then the literature may also generate estimates of the impact of multiple

victimizations. This would have likely added relevant and interesting information to this study. For example, the control site generated a subject charged with three counts of rape in a single arrest. This person was coded the same as a person charged with a single act of rape. This occurred because there was a lack of specificity in the data to decipher how many victims were involved, or how many victimizations occurred. Even if I had been able to decipher the criminal event in such detail, the literature has not developed a framework for valuing multiple victims / victimizations. The paper is limited, then, in that it does not analyze recidivism in sufficient detail to account for and assess the greater number of charges generated by the comparison site. It is important to note, of course, that no study to date has examined more than the controlling offense (the approach I used). I addition, this dissertation is the first paper I am aware of bringing this potential problem to light.

A third limitation was the necessity of restricting the analysis to crime, rather than including other costly and/or beneficial social behavior. The literature which has addressed dependant variables within criminological interventions has shown that crime-related costs overwhelm other potential behaviors. Thus, the omission likely did not impact the overall findings. But the inclusion of other information (i.e. employment and wage info, physical or mental health, mortality, etc.) would have been useful and interesting at least for diagnostic reasons. Inasmuch as the employability, physical and mental health, etc. are important goals for the programs in question, then it would be preferable to provide feedback to policymakers regarding the success of the boot camp at impacting each, in addition to incorporating each into the cost-benefit computations. The research team wanted to collect this type of data from self-report interview with parolees,

and continues to pursue this data. However, this later data has not yet materialized. Thus, the lack of information on other social behavior remains a limitation to the current design.

Gaps in the literature filled

Regardless, this research allowed a strong evaluation of the costs and benefits of the policy choice facing Maryland. Despite limitations, this paper has attempted to be more thorough and complete than most program evaluations in terms of valuation and sensitivity analyses, as well as data collected. Several contributions to the Criminological and Cost-Benefit literature are worth particular mention.

First, prior cost-benefit studies examining prison programs have neglected the income generated by various programs. This is a fairly regular pattern, in fact, in prison management. For example, the budgetary data reported to the state legislature in Maryland reported a "day cost" for facility operations that were derived from costs only. In this raw tables, I found footnotes that presented income information, sometimes in cryptic language and certainly not used in generating the financial summary information for the state. After a rigorous pursuit of the meaning of these lines of financial information, and discussions with multiple state budgetary experts, I was finally able to derive their meaning, the use they should have for the state, and the use they must have for me and this project. However, my identification of these financial lines, interpretation of them, and consolidation of them into the process of summarizing operating information for the two prisons studies was new to the academic literature as well as the state policy makers. As I have argued in this paper, this should not be the case for the

question I address in this document. The department of Corrections will pay for the net cost of a program, not the total expense line of a budget.

Second, this work has introduced and explored the impact of assumptions regarding legal innocence on recidivism costs. This issue has not yet been addressed in the literature, but is endemic to criminological evaluation regardless. Past work drawing on official data has assumed all costs observed were 'fair game' for comparison purposes. At best, they made decision on which costs to include based on availability of data, but not with an explicit understanding or discussion of the implications of innocence to those costs. In contrast, I have approached the issue of innocence in a more tentative manner, presenting low and high estimates of program impacts under extreme versions of the assumption. Although the assumption of guilt is probably more correct than the assumption of 'innocence unless legal guilt is determined,' placing the assumption 'up front' provides interesting information and a clearer picture of sensitivity.

Third, this paper has introduced the idea of accounting for variation in prison costs across security levels within cost-benefit studies. In prior research, these costs were assumed constant across all incarcerated subjects; an assumption that I find troubling. The cost of incarceration varied substantially between facilities in the state of Maryland. By using sentence length and offense severity, I estimated the security level of re-incarcerated subjects. This added a finer grain of measurement with respect to costs endured by the state. As reincarceration is by far the most expensive element of *criminal justice* costs, the added specificity here may play an important role in future cost benefit research, especially when studies are limited to a Criminal Justice model.

A fourth strength of this research was that I used two official data sets to cross-reference information on recidivism. People in the field are often emphatic regarding the poor quality of official criminal record data bases. However, the rate of error is far smaller when two data bases are cross-referenced. The detailed analysis of the error rate in the premier criminal history database in the state of Maryland (CJIS) should demonstrate to the broader research field the necessity of testing the veracity of the 'best' data sources, and leaning toward cross-referencing datasets.

Future Research

Drawing both on the strengths and weaknesses of this dissertation, several avenues for future research are suggested. First, future research may benefit from replicating this design across facilities that vary in their character, program content, or subjects served. This research cannot speak to the broader national debate regarding boot camps in any definitive sense. The study was designed to answer a specific policy question and is tied to the nuances of the two programs compared; it included a limited representation of both prisons and boot camps. That being said, it could contribute to the broader debate by representing one data point among many if the research design were replicated across numerous types of facilities and populations.

Second, the literature would certainly benefit from including the data noted above as missing from this analysis. This includes social behavior other than crime (i.e. mortality, health, education, family functioning, employment, etc.). But it also refers to the specificity of data with regard to crime. This would include public defender usage, intermittent detention, and perhaps detailed information on the actual number and type of

hearings encountered by each recidivist. It is particularly important that analysts collect information at the police-report level so that the number of victims and victimizations might be identified. To facilitate the interpretation of this data, however, future work needs to develop the valuation literature with respect to the costs of multiple victimizations. This is true both of multiple victim events, and especially multiple victimizations. For example, are the costs of being victimized repeatedly linear? Is that 'linearity' similar or different across crime types? Across cost-benefit models?

Third, in reading the valuation literature of 'criminal justice costs' in conjunction with the literature on 'victim costs,' one cannot help but be struck by the difference in methodological creativity and rigor between the two topical areas. The 'victim' literature is far more developed in terms of ideas, methods, creativity, and reliability. Future research would benefit greatly from examining the criminal justice costs methodologies, evoking more creative methods to extrapolate what crime actually costs law enforcement workers and budgets.

For example, valuators of police losses may want to look at the marginal increase in injury for police officers on routine patrol, versus responding to calls for service (by crime type). By combining this information with material and human injury endured by officers along with public health methodologies (i.e. valuating those injuries), the policing literature may realize dramatic changes in the way policing is viewed and valued. For example, the value currently attached to a traffic stop is negligible; it has so little value that most studies ignore it completely or hold it at zero value. However, it is important to note that the traffic stop is one of the leading causes of officer fatality. Likewise, the danger for police officers responding to domestic violence cases is

exceptionally high, relative to other forms of assault calls. In both cases, attaching the injury value to crime types would greatly increase the accuracy of current estimates of police-costs of crime, and bring them into a more accurate realm of valuation.

Fourth, future research may benefit from considering the victimization of research subjects while detained in programs; particularly if the interventions are prison-based. I am unaware of any study to date which assigns value to inmate victimization; I believe this should be addressed. In methodological terms, the idea would represent a creative and interesting elaboration of the cost-benefit methodology, generating more finely grained information that is theoretically relevant. For example, if two prison programs cost exactly the same in all other respects, but differ in terms of violence against inmates while being treated, this information would be valuable in making a policy decision. The issue of whether inmates have 'standing' as citizens is a subject of vigorous debate in the policy arena. However, even skeptics of humanitarian arguments would be hard-pressed to deny that victimization among the confined can generate relevant costs (in terms of risk of retaliation, riot, suicide, later offending, trauma which requires treatment, etc.). Thus, even divorced from the broader debate regarding the "standing" of inmates, inmate experiences are worth valuing and adding into our evaluations. However, the current study does not include these costs because data were not made available with regard to prison victimization.

Finally, future research would benefit from investigating interactions of program utility and subject characteristics amongst these data, as suggested by Uggen (2000). For example, Uggen (2000) found that an age-interaction with regard to a work program offered to parolees. Subjects over age 26 demonstrated significant reductions in

offending if randomly assigned to the intervention, while younger subjects showed no treatment benefits. The literature would benefit from replicating Uggen's work within the context of other intervention types, including boot camp programs for adults. Likewise, the literature would benefit from developing greater theoretical frameworks for understanding and predicting age interactions, as well as interactions with other characteristics of subjects to be served. This would be valuable in terms of policy, as it may lead to a far more efficient deployment of resources. It would also benefit policy in terms of encouraging the development of more nuanced understanding of how and when programs will work for different offenders.

APPENDIX: Figure and Tables

Figure 1. Taxonomy for modeling criminal justice cost of crime: Hypothetical "robbery"

	Primary	**Secondary**
Tangible	"A" Vic. Money Lost Vic. Work Missed Vic. Phys. Injury	"B" Police Expense Court Expense Prison Expense
Intangible	"C" Vic. Fear, Vic. Pain & Suffering	"D" Area Fear of Crime Area 'Flight'

Figure 2. Censoring: *Actual* sentence length equal (hypothetical example)

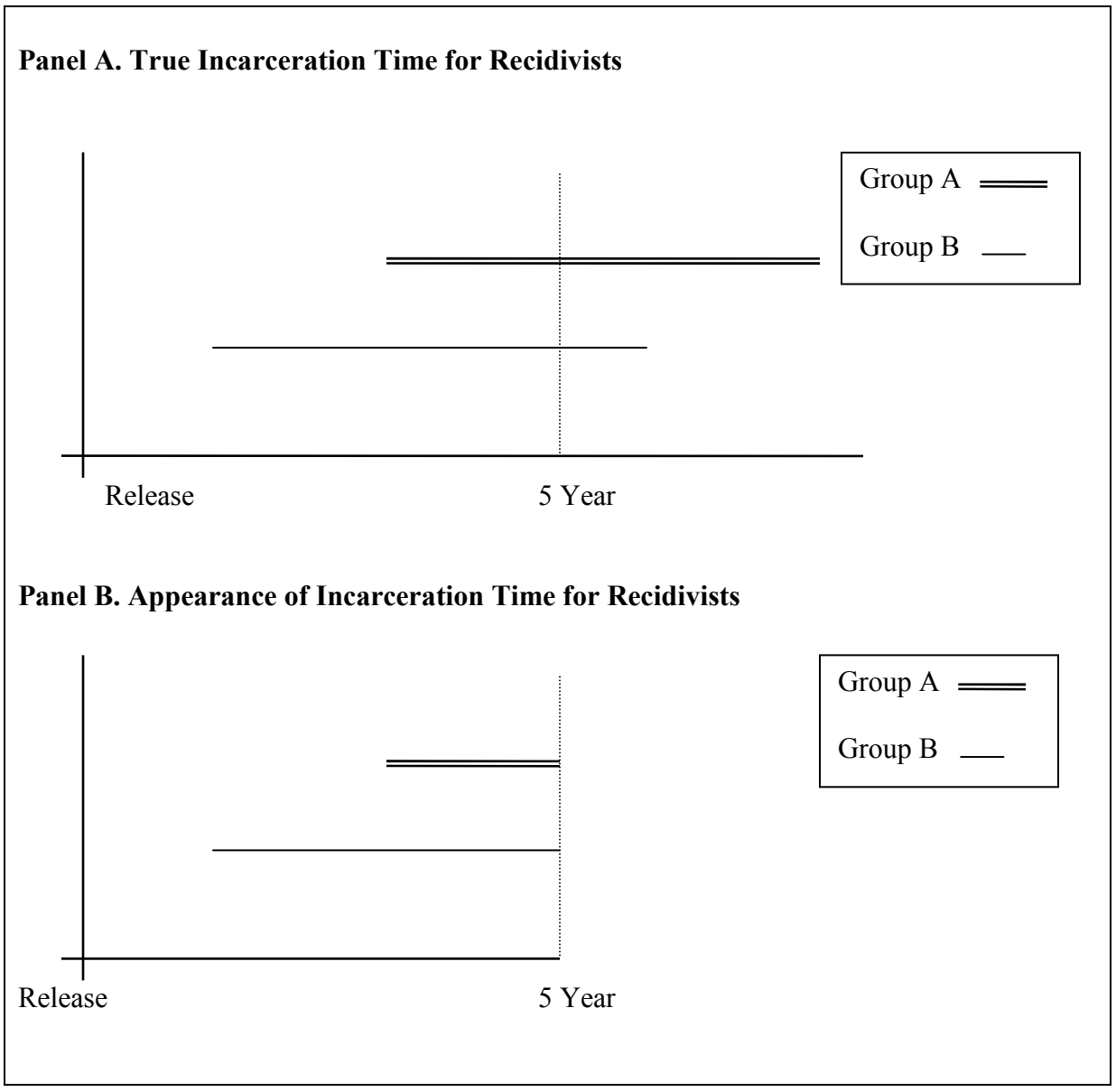

Table 1. The cost of making an arrest: Summary of literature on policing and emergency response (2005 dollars)

	Homicide	Rape	Arson	Rob	Rob (Att)	Agg Assault	Assault (att)	Burgle	Theft	MVT	Forge	Drug Deal	Drug Use	Gamble	DWI	Vandal	Traffic	Dis. Conduct
TOP-DOWN																		
Aos et al. 2001	$16,084	$16,084	---	$16,084	$16,084	$16,084	$16,084	$2,422	$2,422	$2,422	$2,422	$2,422	$2,422	$979	---	---	---	---
Rajkumar et al. 1997	---	---	---	$4,701	$4,701	$4,527	$4,527	$1,276	$805	$935	$611	$25	$25	$8	---	---	---	---
Dade County 1987	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820	$820
BOTTOM-UP																		
Barnet 1985	$213	$213	$213	$213	$213	$213	$213	$213	$213	$213	$213	$213	$213	$56	$56	$56	$56	$56
Austin 1986	$306	$259	$75	$171	$171	$143	$143	$173	$106	$122	---	$247	$247	---	---	$75	$57	$57
Crumpton et al. 2004	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581	$581
Cohen et al. 1994	$198	$134	---	$86	---	$52	$52	---	---	---	---	---	---	---	---	---	---	---
Miller et al. 1996	$1,810	$90	$1,352	$487	$335	$238	$236	$332	$335	$223	---	---	---	---	$54	---	---	---
Cohen 1998	$10,614	$4,636	---	$869	---	$354	---	$747	$264	$756	---	---	---	---	---	---	---	---

Table 2. Victim losses: A comparison of estimates in the literature (2005 values)*

	Homicide	Homicide (att)	Arson	Arson (att)	Rape (adult)	Agg Assault	Assault (att)	Rob	Rob (att)	Burgle	Theft	MVT	DWI	Vandal
Miller et al. 1996*														
Total	$3,643,988	$29,779	$249,240	$19,220	$107,214	$29,780	$2,270	$23,362	$2,344	$1,575	$360	$4,538	$3,332	---
Tangible	$1,275,588	$5,848	$59,520	$18,600	$6,278	$5,848	$162	$6,250	$732	$1,203	$360	$4,166	$1,596	---
Intangible	$2,368,400	$23,932	$189,720	$620	$100,936	$23,932	$2,108	$17,112	$1,612	$372	$0	$372	$1,736	---
NAS 1994														
Total	---	---	---	---	$92,987	$28,383	---	$32,978	---	---	---	---	---	---
Tangible	---	---	---	---	$10,812	$1,487	---	$3,244	---	---	---	---	---	---
Intangible	---	---	---	---	$82,175	$26,896	---	$29,734	---	---	---	---	---	---
Cohen 1988														
Total	---	---	---	---	$92,582	$21,760	---	$22,841	---	$324	$5,677	---	---	---
Tangible	---	---	---	---	$8,380	$770	---	$2,027	---	$324	$5,541	---	---	---
Intangible	---	---	---	---	$84,202	$20,990	---	$20,814	---	$0	$135	---	---	---
McCollister 2006														
Total	$8,747,692	---	$8,657	---	$206,038	$115,155	---	$47,879	---	$4,093	$1,384	---	---	$462
Tangible	$1,174,120	---	$6,455	---	$26,733	$19,157	---	$28,588	---	$352	$12	---	---	$0
Intangible	$7,573,572	---	$2,411	---	$179,455	$104,376	---	$21,517	---	$3,773	$1,373	---	---	$462
Rajkumar & French 1997														
Total	---	---	---	---	---	$46,713	---	$17,713	---	---	---	---	---	---
Tangible	---	---	---	---	---	$305	---	$300	---	---	---	---	---	---
Intangible	---	---	---	---	---	$46,408	---	$17,413	---	---	---	---	---	---

* Note, categories found in Table 1, but missing here, are excluded because they are assumed to have zero victim impacts (i.e. drug offenses).

Table 3. Performance measures: Facility characteristics (2001 – 2003)

	Boot Camp	Control
2001		
Average Pop. of Inmates	312	1,631
Number Staff (total)	105	491
Number Staff (custodial)	84	388
Number Staff (other)	21	103
Inmate-staff ratio (all-staff)	2.97 : 1.00	3.32 : 1.00
Inmate-staff ratio (custodial)	3.71 : 1.00	4.20 : 1.00
2002		
Average Pop. of Inmates	315	1,698
Number Staff (total)	109	494
Number Staff (custodial)	86	407
Number Staff (other)	23	87
Inmate-staff ratio (all-staff)	2.89 : 1.00	3.44 : 1.00
Inmate-staff ratio (custodial)	3.66 : 1.00	4.17 : 1.00
2003		
Average Pop. Of Inmates	350	1,698
Number Staff (total)	102	488
Number Staff (custodial)	80	402
Number Staff (other)	22	86
Inmate-staff ratio (all-staff)	3.43 : 1.00	3.48 : 1.00
Inmate-staff ratio (custodial)	4.38 : 1.00	4.22 : 1.00

Table 4. Group sample size, randomization, and eligibility

	Boot Camp	Control	Total
Cases in study (Count)	**111**	**123**	**234**
Excluded from recidivism study (Count)			
Deported from US	1	0	1
Not released (before Nov. 2006)	2	0	2
Free less than 12 months	3	2	5
Total Excluded	***6***	***2***	***8***
Total available for analysis	***105***	***121***	***226***

Table 5. Annual program budget for 2001 (Converted to 2005 dollars)

	Boot Camp (annual)	Boot Camp (per inmate)	Control (annual)	Control (per inmate)	Difference* BC - Control (per inmate)
Budget Summary					
General Administration	$190,776	$611	$1,493,334	$916	-$304
Custodial Care	$5,335,337	$17,100	$23,141,821	$14,189	$2,912
Dietary Services	$491,037	$1,574	$3,851,367	$2,361	-$788
Plant Operations	$380,023	$1,218	$3,666,919	$2,248	-$1,030
Clinical / Hospital	$918,494	$2,944	$4,978,351	$3,052	-$108
Classification/Rec/Relig	$324,667	$1,041	$2,654,439	$1,627	-$587
Substance Abuse	$107,341	$344	---	---	$344
Total	$7,747,674	$24,832	$39,786,232	$24,394	$439
Appropriation Statement					
Salaries (fringe, etc.)	$5,553,307	$17,799	$26,499,863	$16,248	$1,551
Technical & Special Fees	$60,596	$194	$13,265	$8	$186
Communication	$14,608	$47	$313,686	$192	-$146
Travel	$292	$1	$21,223	$13	-$12
Fuel & Utilities	$172,505	$553	$2,201,214	$1,350	-$797
Vehicles	$80,368	$258	$62,289	$38	$219
Contract services	$1,049,379	$3,363	$8,722,951	$5,348	-$1,985
Supplies and materials	$543,613	$1,742	$765,158	$469	$1,273
Equipment	$4,118	$13	$49,987	$31	-$17
Grants, subsidies, etc	$268,888	$862	$1,079,907	$662	$200
Fixed Charges	$0	$0	$56,691	$35	-$35
Total Operating	$2,133,771	$6,839	$13,273,104	$8,138	-$1,299
Total Expenditures	$7,747,674	$24,832	$39,786,232	$24,394	$439
Special Income					
DOT -Highway reimbursement	*-$727,857*	*-$2,333*	*$0*	*$0*	-$2,333
Inmate welfare fund	*-$255,988*	*-$820*	*-$904,030*	*-$554*	-$266
Net Operating Expenses	**$6,763,829**	**$21,679**	**$38,882,202**	**$23,839**	**-$2,161**

* Note, a negative value indicates the boot camp cost less than the control facility.

Table 6. Annual Program Budget for 2002 (Converted to 2005 dollars)

	Boot Camp (annual)	Boot Camp (per inmate)	Control (annual)	Control (per inmate)	Difference* BC - Control (per inmate)
Budget Summary					
General Administration	$369,374	$1,173	$4,717,629	$2,778	-$1,606
Custodial Care	$4,932,647	$15,659	$23,972,660	$14,118	$1,541
Dietary Services	$801,223	$2,544	$4,151,443	$2,445	$99
Plant Operations	$464,285	$1,474	$3,080,511	$1,814	-$340
Clinical / Hospital	$1,049,100	$3,330	$3,454,167	$2,034	$1,296
Classification/Rec/Relig	$419,117	$1,331	$2,135,959	$1,258	$73
Substance Abuse	$53,019	$168	---	---	$168
Total	$8,088,765	$25,679	$41,512,368	$24,448	$1,231
Appropriation Statement					
Salaries (fringe, etc.)	$5,679,815	$18,031	$29,936,437	$17,630	$401
Technical & Special Fees	$40,289	$128	$12,434	$7	$121
Communication	$21,690	$69	$205,940	$121	-$52
Travel	$169	$1	$18,675	$11	-$10
Fuel & Utilities	$191,818	$609	$1,823,920	$1,074	-$465
Vehicles	$74,179	$235	$66,304	$39	$196
Contract services	$1,205,948	$3,828	$7,417,938	$4,369	-$540
Supplies and materials	$619,808	$1,968	$854,366	$503	$1,464
Equipment	$10,838	$34	$86,072	$51	-$16
Grants, subsidies, etc	$243,635	$773	$1,041,326	$613	$160
Fixed Charges	$575	$2	$48,956	$29	-$27
Total Operating	$2,368,661	$7,520	$11,563,497	$6,810	$709
Total Expenditures	$8,088,765	$25,679	$41,512,368	$24,448	$1,231
Special Income					
DOT -Highway reimbursement	-$816,096	-$2,591	$0	$0	-$2,591
Inmate welfare fund	-$208,389	-$662	-$909,972	-$536	-$126
Net Operating Expenses	**$7,064,280**	**$22,426**	**$40,602,396**	**$23,912**	**-$1,486**

Table 7. Annual Program Budget for 2003 (Converted to 2005 dollars)

	Boot Camp (annual)	Boot Camp (per inmate)	Control (annual)	Control (per inmate)	Difference* BC - Control (per inmate)
Budget Summary					
General Administration	$358,474	$1,024	$5,723,569	$3,395	-$2,371
Custodial Care	$5,263,506	$15,039	$23,128,217	$13,718	$1,321
Dietary Services	$825,740	$2,359	$4,013,199	$2,380	-$21
Plant Operations	$425,257	$1,215	$3,638,966	$2,158	-$943
Clinical / Hospital	$1,026,995	$2,934	$4,191,792	$2,486	$448
Classification/Rec/Relig	$411,128	$1,175	$2,330,001	$1,382	-$207
Substance Abuse	$132,186	$378	$0	$0	$378
Total	$8,443,286	$24,124	$43,025,744	$25,519	-$1,396
Appropriation Statement					
Salaries (fringe, etc.)	$6,021,268	$17,204	$30,349,707	$18,001	-$797
Technical & Special Fees	$7,228	$21	$10,393	$6	$14
Communication	$18,731	$54	$231,616	$137	-$84
Travel	$212	$1	$17,811	$11	-$10
Fuel & Utilities	$185,352	$530	$2,313,401	$1,372	-$843
Vehicles	$60,003	$171	$93,203	$55	$116
Contract services	$1,216,545	$3,476	$8,074,823	$4,789	-$1,313
Supplies and materials	$612,785	$1,751	$800,044	$475	$1,276
Equipment	$6,350	$18	$32,047	$19	-$1
Grants, subsidies, etc	$314,812	$899	$1,032,882	$613	$287
Fixed Charges	$0	$0	$69,818	$41	-$41
Total Operating	$2,414,790	$6,899	$12,665,644	$7,512	-$613
Total Expenditures	$8,443,286	$24,124	$43,025,744	$25,519	-$1,396
Special Income					
DOT -Highway reimbursement	*-$601,714*	*-$1,719*	*$0*	*$0*	*-$1,719*
Inmate welfare fund	*-$258,692*	*-$739*	*-$880,277*	*-$522*	*-$217*
State Alien Assistance Program	*$0*	*$0*	*-$474,355*	*-$281*	*$281*
Military Operation Maintenance	*-$541*	*-$2*	*-$1,987*	*-$1*	*$0*
Net Operating Expenses	**$7,582,339**	**$21,664**	**$41,669,126**	**$24,715**	**-$3,051**

* Note, a negative value indicates the boot camp cost less than the control facility.

Table 8. Valuation categories: Charges contained in each crime-type category for cost-assessment

Valuated Category	***Actual Charges***
Homicide	*Homicide-1st, Homicide-2nd*
Homicide (attempt)	*Att-Homicide-1st, Att-Homicide-2nd*
Rape	*Rape-1, Rape-2, Sex off-1, sex off-2, sex off-3*
Arson	*Arson 1*
Arson (attempt)	*Att-arson*
Robbery	*Robbery, armed robbery, car-jacking*
Robbery (attempt)	*Att-robbery, conspire-robbery*
Agg. Assault	*Assault-1, Assault-2*
Assault (attempt)	*Resisting Arrest*
Burglary	*Burglary-1, Burglary-2, Att Burglary*
Larceny / Forgery	*Theft, forgery*
MV Theft	*Motor Vehicle Theft*
Drug Dealing	*Possession with intent to distribute, distribution*
Drug Use	*Possession*
Gambling	*Gambling, conspire to Gamble*
DWI	*Driving while impaired*
Vandalism	*Vandalism, Malicious Destruction*
Traffic	*Reckless Driving, Driving without License, etc.*
Disorderly conduct	*Disorderly Conduct, Vagrancy, Rogue and Vagabond*
Weapon / Gun	*Possession of Firearm, Firearm in vehicle, use in felony, weapon other*

Table 9. Cost of making an arrest: Values used in criminal justice computations

	Lowest	Highest	Best*	Best (Source)
Homicide	$198	$16,084	$10,614	Cohen 1998
Homicide (attempt)	$52	$16,084	$354	Cohen 1998**
Rape	$90	$16,084	$4,636	Cohen 1998
Arson	$75	$1,352	$1,352	Miller et al. 1996
Arson (attempt)	$75	$1,352	$1,352	Miller et al. 1996
Robbery	$86	$16,084	$869	Cohen 1998
Robbery (attempt)	$86	$16,084	$335	Miller et al. 1996
Agg. Assault	$52	$16,084	$354	Cohen 1998
Assault (attempt)	$52	$16,084	$236	Miller et al. 1996
Burglary	$173	$2,422	$747	Cohen 1998
Larceny / Forgery	$106	$2,422	$264	Cohen 1998
MV Theft	$122	$2,422	$756	Cohen 1998
Drug Dealing	$25	$2,422	$25	Rajkumar & French 1997
Drug Use	$25	$2,422	$25	Rajkumar & French 1997
Gambling	$8	$979	$20	Minimal Police Time***
DUI / DWI	$54	$820	$54	Miller et al. 1996
Vandalism	$56	$820	$75	Austin 1986
Traffic	$56	$820	$57	Austin 1986
Disorderly conduct	$56	$820	$57	Austin 1986
Weapon / Gun	$20	$20	$20	Minimal Police Time***

* 'Best' estimates chosen based on bottom-up methodologies, age of data, and completeness of data sources

** No valuation exists for attempted murder, so 'injurious assault' substituted.

*** One hour of entry-level police resources for state of Maryland; $17 / hr, + 25% fringe, = $20.25 / hour

Table 10. Cost of prosecution: Values used in criminal justice computations (2005 values)

	Cost (average)
Hearing Type	
Arraignment	$653
Serious Hearing	$1,641
Minor Hearing	$560
Sentencing Hearing	$909
Total Cost per Outcome	
Dismissed	$653
Nolle Prossed or STET	$1,213
Not Guilty	$2,854
Conviction	$3,763

Table 11. Net-cost of incarceration: Average day cost across Maryland facilities (2005)

Facility	Facility (Full Name)	Security	Day Cost per Inmate (net)
Jail / Pre-release			
BCDC	Baltimore City Detention Center	Pre-release	$60.40
EPRU	Eastern Pre-Release Unit	Pre-release	$50.65
JPRU	Jessup Pre-Release Unit	Pre-release	$52.08
PHPRU	Poplar Hill Pre-Release Unit	Pre-release	$46.05
SMPRU	Southern Maryland Pre-Release Unit	Pre-release	$50.01
BCCC	Baltimore City Correctional Center	Pre-release	$48.53
BPRU	Baltimore Pre-Release Unit	Pre-release	$45.55
Average (Jail)			**$50.47**
Minimum			
BCF	Brockbridge Correctional Facility	Minimal	$54.97
CLF	Central Laundry Facility	Minimal	$54.47
MTC *	Metropolitan Transition Center	Minimal	$65.38
Average (Minimal)			**$58.27**
Medium			
ECI	Eastern Correctional Institution	Medium	$58.21
MCI-H	Maryland Correctional Institution - Hagerstown	Medium	$60.14
MCI-J	Maryland Correctional Institution - Jessup	Medium	$62.87
MCTC	Maryland Correctional Training Center	Medium	$42.58
NBCI	North Branch Correctional Institution	Medium	$68.04
RCI	Roxbury Correctional Institution	Medium	$48.09
WCI	Western Correctional Institution	Medium	$63.01
Average (Medium)			**$57.56**
Maximum			
MHC	Maryland House of Correction	Maximum	$74.82
JCI	Jessup Correctional Institution (formerly MHC Annex)	Maximum	$77.48
Average (Maximum)			**$76.15**

Note: list omits facilities that are group homes, female facilities, or diagnostic centers that are substntively transient. Eastern-annex is averaged within ECI.

* MTC, the control site in this study, contains inmates who would be considered maximum, medium, and minimum security. However, they are all in the last two years of their respective sentences if in this facility. Substantively, a person must be sentenced as a minimum security offender in order to qualify for a full term assignment here, so it is designated as "minimum security."

Table 12. Victim losses: 'Best' estimates of tangible and intangible values from literature

	Intangible	Tangible	Total	Source
Violent				
Homicide	$2,368,400	$1,275,588	$3,643,988	Miller et al. 1996
Homicide (attempt)*	$23,932	$5,848	$29,780	Miller et al. 1996
Rape	$100,936	$6,278	$107,214	Miller et al. 1996
Arson	$189,720	$59,520	$249,240	Miller et al. 1996
Arson (attempt)	$620	$18,600	$19,220	Miller et al. 1996
Robbery	$17,112	$6,250	$23,362	Miller et al. 1996
Robbery (attempt)	$1,612	$732	$2,344	Miller et al. 1996
Assault	$23,932	$5,848	$29,780	Miller et al. 1996
Assault (attempt)	$2,108	$162	$2,270	Miller et al. 1996
Disorderly conduct	$0	$162	$162	Miller et al. 1996 *
Property				
Burglary	$372	$1,203	$1,575	Miller et al. 1996
Larceny / Forgery	$0	$360	$360	Miller et al. 1996
MV Theft	$372	$4,166	$4,538	Miller et al. 1996
Vandalism	$0	$449	$449	McCollister 2006
Victimless				
DWI	$1,736	$1,596	$3,332	Miller et al. 1996
Drug Dealing	$0	$0	$0	Assumes no Victim Cost
Drug Use	$0	$0	$0	Assumes no Victim Cost
Gambling	$0	$0	$0	Assumes no Victim Cost
Traffic	$0	$0	$0	Assumes no Victim Cost
Gun / Weapon	$0	$0	$0	Assumes no Victim Cost

* Values of 'non injury assault' substituted assume no intangible loss.

Table 13. Descriptive statistics across facilities (N = 226)

	Boot Camp N=105	**Control N=121**
Age M (SD)	23.11 (4.30)	23.46 (3.97)
Race: African American (%)	83%	82%
Prior arrests M (SD)	5.27 (4.29)	5.32 (4.38)
Prior convictions M (SD)	2.54 (1.85)	2.52 (1.92)
Offense of conviction (%)		
Violent	2%	2%
Drug	90%	93%
Property and other	8%	5%
Drop outs (%)	15%	11%
Returned to Baltimore City (%)	54%	61%
Intensive parole (%)	50%	31% **

* * Difference between boot camp and comparison is significant, p<.01

Table 14. Descriptive statistics of basic recidivism information for both groups

	Boot Camp	Comparison
Re-arrested (count, %)	63 (60%)	81 (67%)
Re-arrest events (count)	114	186
Reconviction events (count)	39	63
Re-incarceration events: *minimal* time assumed (count)	31	45
Re-incarceration events: *maximum* time assumed (count)	35	50

Table 15. Average days served per facility (2001 – 2003)

	Boot Camp	Control	Total
Released Late (Count)			
LT 30 days	3	24	27
GT 30 days	13	16	29
Total	*16*	*40*	*56*
Released Early (Count)			
LT 30 days	*8*	*8*	*16*
GT 30 Days	*5*	*3*	*8*
Total	*13*	*11*	*24*
Average Time Served (Days)			
Among Late Releases (M, SD)	416.91 (244.33)	270.97 (180.34)	323.61 (215.60)
Among Early Releases (M, SD)	140.23 (55.29)	156.36 (43.09)	147.62 (49.87)
Total Population (M, SD)	***208.34 (100.14)***	***196.21 (54.09)***	***201.85 (78.95)***

Table 16. Net housing costs for each inmate (per day): 2005 dollars

Year	Boot Camp	Control	Difference BC - Control
2001	$59.36	$65.28	-$5.92
2002	$61.41	$65.48	-$4.07
2003	$59.32	$67.67	-$8.35
Average	**$60.03**	**$66.14**	**-$6.11**

Table 17. Average program costs: Net cost estimates (2005 dollars)

	Boot Camp	Control	Difference BC - Control
Corrected Cost Estimates			
Avg. Inmate cost / day	$60.03	$66.14	
Average days per served	*208.34 (100.14)*	*196.21 (54.09)*	
Average cost per sentence	$12,506.65	$12,977.33	-$470.68

Table 18. Program costs, accounting for value of lost inmate labor

	Boot Camp	Control	Difference BC - Control
Corrected Cost Estimates			
Inmate cost per day	$12,507	$12,977	
Lost labor per day	$9,011	$8,486	
Average cost per sentence	*$21,518*	*$21,463*	$55

Table19. Offenses *charged* in recidivism data set (N = 226 potential offenders)

Charge	Boot Camp (count)	Boot Camp (per inmate)	Control (count)	Control (per inmate)	Difference BC - Control (per inmate)
Homicide	0	0.00	3	0.02	-0.02
Homicide (attempt)	4	0.04	4	0.03	0.01
Arson	2	0.02	0	0.00	0.02
Assault (major or minor)	35	0.33	55	0.45	-0.12
Firearm	31	0.30	31	0.26	0.04
Kidnapping	0	0.00	1	0.01	-0.01
Robbery	5	0.05	16	0.13	-0.08
Sex offense	0	0.00	3	0.02	-0.02
Stalking	1	0.01	0	0.00	0.01
Resist / Obstruct	15	0.14	23	0.19	-0.05
Weapon -	1	0.01	3	0.02	-0.01
Total Violent	**94**	**0.90**	**139**	**1.15**	**-0.25**
					0.00
Alcohol	0	0.00	3	0.02	-0.02
Drug Possess - Hard	16	0.15	38	0.31	-0.16
Drug Possess - Unknown	71	0.68	84	0.69	-0.01
Drug Possess - MJ	15	0.14	24	0.20	-0.06
Drug Sales - Unknown	64	0.61	125	1.03	-0.42
Drug Sales - Hard	1	0.01	7	0.06	-0.05
Total Drug	**167**	**1.59**	**281**	**2.32**	**-0.73**
					0.00
Burglary	8	0.08	11	0.09	-0.01
Mal Destruction	5	0.05	14	0.12	-0.07
MVT	13	0.12	12	0.10	0.02
Larceny	2	0.02	3	0.02	0.00
Theft	44	0.42	32	0.26	0.16
Trespass	11	0.10	9	0.07	0.03
Total Property	**83**	**0.79**	**81**	**0.67**	**0.12**
					0.00
Conspiracy	1	0.01	0	0.00	0.01
Disorderly	1	0.01	4	0.03	-0.02
Escape	0	0.00	2	0.02	-0.02
Gambling	1	0.01	2	0.02	-0.01
Hire Minor/ Cont. Delinq.	2	0.02	2	0.02	0.00
Reckless Endangerment	5	0.05	8	0.07	-0.02
Rogue and Vagabond	0	0.00	3	0.02	-0.02
Traffic	7	0.07	22	0.18	-0.11
Total Other	**17**	**0.16**	**43**	**0.36**	**-0.20**
All Offenses	**361**	**3.44**	**544**	**4.50**	**-1.06**

Table 20. Comparison of *controlling* offenses: All arrest events across facilities

	Boot Camp		Control		Difference
	Count	Per Inmate (N = 105)	Count	Per Inmate (n=121)	BC - Control
Homicide	0	0.00	3	0.02	-0.02
Homicide (attempt)	2	0.02	1	0.01	0.01
Sex offense	0	0.00	1	0.01	-0.01
Arson	1	0.01	0	0.00	0.01
Assault (major or minor)	24	0.23	31	0.25	-0.02
Disorderly	2	0.02	11	0.09	-0.07
Total Violent	**29**	**0.28**	**47**	**0.39**	**-0.11**
Drug Posses or Sales	**52**	**0.50**	**96**	**0.80**	**-0.30**
Burglary	3	0.03	3	0.02	0.01
MVT	6	0.06	7	0.06	0.00
Theft	12	0.11	11	0.09	0.02
Vandal	0	0.00	4	0.03	-0.03
Total Property	**21**	**0.20**	**25**	**0.21**	**-0.01**
Gambling	1	0.01	2	0.02	-0.01
Firearm	2	0.02	2	0.02	0.00
Weapon - other	1	0.01	1	0.01	0.00
Other*	5	0.05	5	0.04	0.01
Traffic	3	0.03	8	0.07	-0.04
Total Other	**13**	**0.12**	**18**	**0.15**	**-0.03**
All Offenses	**114**	**1.10**	**186**	**1.54**	**-0.44**

Note, controlling offense refers to the offense which generates the most monetary harm w.r.t. 'full' cost-benefit criteria.

Table 21. Summary of cost elements across facilities & assumptions

	BC (total)	Control (total)	Difference* (BC – Control)
Program Costs			
1. Observed	$12,507	$12,977	-$470
2. Labor-adjusted	$21,518	$21,463	$55
Arrest Costs			
1. Observed	$295	$695	-$400
2. Legal Guilt Only	$98	$219	-$121
Court			
1. Observed	$2,292	$3,149	-$857
2. Legal Guilt Only	$1,434	$1,965	-$531
Incarceration (by security level)			
1. Minimum	$24,005	$22,270	$1,735
2. Maximum	$37,204	$40,174	-$2,970
Victim Tangible			
1. Observed	$2,084	$33,439	-$31,355
2. Legal Guilt Only	$310	$10,989	-$10,679
Victim Intangible			
1. Observed	$6,726	$65,374	-$58,648
2. Legal Guilt Only	$979	$21,830	-$20,851

* Note, a negative value indicates the boot camp cost less than the control site.

Table 22. Court outcomes of recidivism arrests: Comparing facilities

	Cost per case	Boot Camp			Control		
		Count	%	Cost	Count	%	Cost
Guilty	$3,763	40	*35.1*	$150,520	63	*33.9*	$237,069
Not Guilty	$2,854	4	*3.5*	$11,416	6	*3.2*	$17,124
Nolle Pros/ Stet	$1,213	59	*51.8*	$71,567	90	*48.4*	$109,170
Dismissed / PJ	$653	11	*9.6*	$7,183	27	*14.5*	$17,631
Total per facility		114	*100.0*	$240,686	186	*100.0*	$380,994
Total per inmate served				**$2,292.25**			**$3,148.71**

Note, court outcome refers to the controlling offense; the most serious outcome of all charges involved in the arrest event.

Table 23. Cost-benefit models: Best models and sensitivity to 'legal guilt' assumption

	Boot Camp (total)*	Control (total)*	Difference** (BC – Control)
Criminal Justice			
Observed costs	$52,298	$56,995	-$4,697
Legal guilt only	$51,243	$55,335	-$4,092
Direct Loss			
Observed costs	$54,382	$90,434	-$36,052
Legal guilt only	$51,553	$66,324	-$14,771
Full Cost-Benefit			
Observed costs	$61,108	$155,808	-$94,700
Legal guilt only	$52,532	$88,154	-$35,622

* Note, computations use costs derived from assumption that maximum sentences will be served.
** A negative value for 'Difference' column indicates the boot camp cost less than the control site.

REFERENCES

Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. *Criminology, 30*(1), 47-87.

Agnew, R. (1999). A general strain theory of community differences in crime rates. *Journal of Research in Crime and Delinquency, 36*(2), 123.

Agnew, R. (2001). Building on the foundation of general strain theory: Specifying the types of strain most likely to lead to crime and delinquency. *Journal of Research in Crime and Delinquency 38*(4), 319-361.

Anderson, E. (1999). *Code of the street: Decency, violence, and the moral life of the inner city*. New York, NY: W.W. Norton.

Aos, S., Phipps, P., Barnoski, R., and Lieb, R. (2001a). The comparative costs and benefits of programs to reduce crime: A review of research findings with implications for Washington State. In B. Welsh, Farrington, D., and Sherman, L. (Ed.), *Costs and benefits of preventing crime*. Boulder, CO: Westview Press.

Aos, S., Phipps, P., Barnoski, R., and Lieb, R. (2001b). The comparative costs and benefits of programs to reduce crimes: Version 4.0. Olympia, WA: Washington State Institute for Public Policy.

Armstrong, G. S. (2004). Boot camps as a correctional option. In D. Mackenzie, and Armstrong, G.S. (Ed.), *Correctional boot camps: Military basic training or a model for corrections?* Thousand Oaks, CA: Sage.

Armstrong, G. S., & MacKenzie, D. L. (2003). Private versus public juvenile correctional facilities: Do differences in environmental quality exist? *Crime & Delinquency, 49*(4), 542.

Austin, J. (1986). Using early release to relieve prison crowding: A dilemma in public policy. *Crime & Delinquency, 32*, 404.

Barnett, W. (1985). *The Perry Preschool Program and its long-term effects: A benefit-cost analysis*. Ypsilanti, MI: High/Scope Educational Research Foundation.

Barnett, W. S. (1992). Benefits of compensatory preschool education. *Journal of Human Resources, 27*, 279.

Barnett, W. S. (1993). Benefit-cost analysis of preschool education: Findings from a 25-year follow-up. *American Journal of Orthopsychiatry, 63*, 500.

Barnett, W. S., & Escobar, C. M. (1989). Research on the cost effectiveness of early educational intervention: Implications for research and policy. *American Journal of Community Psychology, 17*, 677.

Bartley, W. (2000). *A valuation of specific crime rates.* Vanderbilt, TN: Dissertation

Blumstein, A., and Beck, A. (2005). Reentry as a transient state between liberty and recommitment. In J. Travis, and Visher, C. (Ed.), *Prisoner reentry and crime in America*. New York, NY: Cambridge University Press.

Boardman, A., Greenberg, D., Vining, A., and Weimer, D. (2001). *Cost-benefit analysis: Concepts and practices*. Upper Saddle River, NJ: Prentice Hall, Inc.

Bottcher, J., & Ezell, M. E. (2005). Examining the effectiveness of boot camps: A randomized experiment with a long-term follow up. *Journal of Research in Crime and Delinquency, 42*(3), 309.

Bourque, B., Cronin, R., and Pearson, F. (1996). Boot camps for juvenile offenders: An implementation evaluation of three demonstration programs. Washington, D.C.: U.S. National Institute of Justice.
Cohen, M. (1988). Pain, suffering, and jury awards: A study of the cost of crime to victims. *Law and Society Review, 22*, 537-555.
Cohen, M. (1998). The monetary value of saving a high-risk youth. *Journal of Quantitative Criminology, 14*(1), 5-33.
Cohen, M. (2001). The crime victim's perspective in cost-benefit analysis: The importance of monetizing tangible and intangible crime costs. In B. Welsh, Farrington, D., and Sherman, L. (Ed.), *Costs and benefits of preventing crime*. Boulder, CO: Westview Press.
Cohen, M., Miller, T., and Rossman, S. (1994). The costs and consequences of violent behavior in the United States. In Roth, A. (Ed.), *Understanding and preventing violence* (Vol. 4, pp. 67-166). Washington, D.C.: National Academy Press.
Cohen, M. A. (2005). *The costs of crime and justice*. New York, NY: Rutledge.
Cohen, M. A., & Miller, T. R. (1998). The cost of mental health care for victims of crime. *Journal of Interpersonal Violence, 13*(1), 93.
Cohen, M. A., Rust, R. T., & Steen, S. (2004). Willingness-to-pay for crime control programs. *Criminology, 42*(1), 89.
Cook, P. J., & Ludwig, J. (2000). *Gun violence: The real costs*. New York, NY: Oxford University Press.
Crumpton, D., Brekhus, J., Weller, J., and Finigan, M. (2004). Cost analysis of Anne Arundel County, Maryland drug court. Portland, OR: NPC Research
Czelen, J., Wheeler, C., and Mott, P. (1978). The cost of justice: An analysis of case processing in the Bronx justice system. New York, NY: Mott-McDonald Associates, Inc.
DiIulio, J. J. (1987). *Governing prisons: A comparative study of correctional management*. New York, NY: Collier Macmillan.
Forst, B. (2004). *Errors of justice: Nature, sources, and remedies*. New York, NY: Cambridge University Press.
French, M., McCollister, K., Alexandre, P., Chitwood, D., and McCoy, C. (2004). Revolving roles in drug-related crime: The cost of chronic drug users as victims and perpetrators. *Journal of Quantitative Criminology, 20*(3).
French, M., Salome, H. J., & Krupski, A. (2000). Benefit-cost analysis of residential and outpatient addiction treatment in the state of Washington. *Evaluation Review, 24*(6), 609.
French, M. T., Salome, H. J., & Carney, M. (2002). Using the DATCAP and ASI to estimate the costs and benefits of residential addiction treatment in the state of Washington. *Social Science & Medicine, 55*(12), 2267.
Goffman, E. (1967). *Interaction ritual; essays on face-to-face behavior*. Garden City, NY: Anchor Books.
Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.
Gover, A. R., Gover, A. R., MacKenzie, D. L., & Styve, G. J. (2000). Boot camps and traditional correctional facilities for juveniles: A comparison of the participants, daily activities, and environments. *Journal of Criminal Justice, 28*(1), 53.

Gover, A. R., Gover, A. R., McKenzie, D. L., & Armstrong, G. S. (2000). Importation and deprivation explanations of juveniles' adjustment to correctional facilities. *International Journal of Offender Therapy and Comparative Criminology, 44*(4), 450.

Gover, A. R., & MacKenzie, D. L. (2003). Child maltreatment and adjustment to juvenile correctional institutions. *Criminal Justice and Behavior, 30*(3), 374.

Greenwood, P., Karoly, L., Everingham, S., Hoube, J., Kilburn, M.R., Rydell, C.P., Sanders, M., and Chiesa, J. (2001). Estimating the costs and benefits of early childhood interventions: Nurse home visits and the Perry Preschool. In B. Welsh, Farrington, D., and Sherman, L. (Ed.), *Costs and benefits of preventing crime*. Boulder, CO: Westview Press.

Hawkins, J. D., Catalano, R. F., Kosterman, R., Abbott, R., & Hill, K. G. (1999). Preventing adolescent health-risk behaviors by strengthening protection during childhood. *Archives Of Pediatrics & Adolescent Medicine, 153*(3), 226-234.

Heckman, J. and Smith, J. (1995). Assessing the case for social experiments. Journal of Economic Perspectives, 9(3), 85-110.

Hoen, J., Berger, M., and Blomquest, G. (1987). A hedonic model of interregional wages, rents, and amenity values. *Journal of Regional Science, 27*(4), 605-620.

Kingsnorth, R. F., MacIntosh, R. C., & Sutherland, S. (2002). Criminal charge or probation violation? Prosecutorial discretion and implications for research in criminal court processing. *Criminology, 40*(3), 553.

Langan, P., and Levin, D. (2002). Recidivism of prisoners released in 1994. Washington, D.C., U.S. Office of Justice Programs.

Larson, C. (1983). Costs of incarceration and alternatives: A case study of burglary. Tempe, AZ: Arizona State University (unpublished).

Logan, T., Hoyt, W., McCollister, K., French, M., Luekfeld, C., and Minton, L. (2004). Economic evaluation of drug court: Methodology, results, and policy implications. *Evaluation and Program Planning, 27*, 381-396.

MacCoun, R., and Reuter, P. (2001). *Drug war heresies: Learning from other vices, times and places*. New York, NY: Cambridge University Press.

Machin, D., Cheung, Y., and Parmer, M. (2006). *Survival analysis: A practical approach* (2 ed.). Hoboken, NJ: Wiley and Sons, Ltd.

Mackenzie, D., & Armstrong, G. S. (2004). *Correctional boot camps: Military basic training or a model for corrections?* Thousand Oaks, CA: Sage.

MacKenzie, D.L., Bierie, D. (2006). Follow Up Study of the Randomized Boot Camp Experiment: Final Report. *Governor's Office of Crime Control and Prevention, state of Maryland.*

MacKenzie, D. L. (2001). Corrections and sentencing in the 21st century: Evidence-based corrections and sentencing. *Prison Journal, 81*(3), 299.

MacKenzie, D. L., Brame, R., & McDowall, D. (1995). Boot camp prisons and recidivism in eight states. *Criminology, 33*, 327.

MacKenzie, D. L., & Piquero, A. (1994). The impact of shock incarceration programs on prison crowding. *Crime & Delinquency, 40*, 222.

MacKenzie, D. L., Wilson, D. B., & Armstrong, G. S. (2001). The impact of boot camps and traditional institutions on juvenile residents: Perceptions, adjustment, and change. *Journal of Research in Crime and Delinquency, 38*(3), 279.

MacKenzie, D. L., Wilson, D. B., & Kider, S. B. (2001). Effects of correctional boot camps on offending. *Annals of the American Academy of Political and Social Science, 578*, 126.

Macmillan, R. (2000). Adolescent victimization and income deficits in adulthood: Rethinking the costs of criminal violence from a life-course perspective. *Criminology, 38*(2), 553.

McCollister, K. E., French, M. T., & Prendergast, M. L. (2004). Long-term cost effectiveness of addiction treatment for criminal offenders. *Justice Quarterly, 21*(3), 659.

Miller, T., Cohen, M., and Weirsema, B. (1996). Victim costs and consequences: A new look. Washington, D.C.: National Institute of Justice.

Mitchell, O., MacKenzie, D. L., Gover, A. R., & Styve, G. J. (2001). The influences of personal background on perceptions of juvenile correctional environments. *Journal of Criminal Justice, 29*(1), 67.

Morash, M., & Rucker, L. (1990). A critical look at the idea of boot camp as a correctional option. *Crime & Delinquency, 36*(2), 204.

Nagin, D. (2001). Measuring economic benefits of developmental prevention programs. In B. Welsh, Farrington, D., and Sherman, L. (Ed.), *Costs and benefits of preventing crime*. Boulder, CO: Westview Press.

Olds, D. L., Henderson, C. R., Jr., & Tatelbaum, R. (1988). Improving the life-course development of socially disadvantaged mothers: A randomized trial of nurse home visitation. *American Journal of Public Health, 78*, 1436.

Pagorsky, G., Piquero, A., and Paternoster, R. (2004). Modeling change in perceptions about sanction threats: The neglected linkage in Deterrence theory. *Journal of Quantitative Criminology, 20*(4), 343 - 369.

Peters, M. (1996). Evaluation of the impact of boot camps for juvenile offenders: Mobile interim report. In Washington, D.C.: U.S. Department of Justice.

Rajkumar, A., and French, M. (1994). Method and calculations for the tangible costs of crime (working paper). Research Triangle Park, NC: Research Triangle Institute.

Rajkumar, A., and French, M. (1997). Drug abuse, crime costs, and the economic benefit of treatment. *Journal of Quantitative Criminology, 13*(3), 291-323.

Rice, D., and MacKenzie, D.E. (1989). Cost of injury in the United States: A report to congress. San Francisco, CA: Institute for Health, Aging, and Injury Prevention Center.

Rosenfeld, R., Wallman, J., Fornango, R. (2005). The contribution of ex-prisoners to crime rates. In J. Travis, and Visher, C. (Ed.), *Prisoner reentry and crime in America*. New York, NY: Cambridge University Press.

Salome, H., French, M., Miller, M., and McClellan, T. (2003). Estimating the client costs of addiction treatment: First findings from the client drug abuse treatment cost analysis program. *Drug and Alcohol Dependence, 71*, 195-206.

Schmidt, P., and Whitte, A. (1988). *Predicting recidivism using survival models*. New York, NY: Springer Verlag.

Schweinhart, L., Barnes, H., and Weikart, D., with Barnet, W., and Epstein, A. (1993). *Significant benefits: The High/Scope Perry Preschool through age 27*. Ypsilanti, MI: The High/Scope Press.

Sherman, L. W. (1993). Defiance, deterrence, and irrelevance: A theory of the criminal sanction. *Journal of Research in Crime and Delinquency, 30*, 445.

Skogan, W. G., & Frydl, K. (2004). *Fairness and effectiveness in policing: The evidence*. Washington, D.C.: National Academies Press.

Thomas, D., and Peters, M. (1996). Evaluation of the impact of boot camps for juvenile offenders: Cleveland interim report. Washington, D.C.: U.S. Department of Justice.

Turner, S., Petersilia, J., & Deschenes, E. P. (1992). Evaluating intensive supervision probation/parole (ISP) for drug offenders. *Crime & Delinquency, 38*, 539.

Uggen, C. (2000). Work as a turning point in the life course of criminals: A duration model of age, employment, and recidivism. *American Sociological Review, 65*(4), 529.

Welsch, B., and Farrington, D. (2001a). Assessing the economic costs and benefits of prevention. In B. Welsh, Farrington, D., and Sherman, L. (Ed.), *Costs and benefits of preventing crime*. Boulder, CO: Westview Press.

Welsch, B., and Farrington, D. (2001b). A review of research on the monetary value of preventing crime. In B. Welsh, Farrington, D., and Sherman, L. (Ed.), *Costs and benefits of preventing crime.* Boulder, CO: Westview Press.

Welsch, B., Farrington, D., and Sherman, L. (2001a). Improving confidence in what works and saves money in preventing crime: Priorities for research. In B. Welsh, Farrington, D., and Sherman, L. (Ed.), *Costs and benefits of preventing crime*. Boulder, CO: Westview Press.

Welsh, B., Farrington, D., and Sherman, L. (2001b). *Costs and benefits of preventing crime*. Boulder, CO: Westview Press.

Zarkin, G., Dunlap, L. Belenko, S., and Dynia, P. (2005). A benefit-cost analysis of the Kings County District Attorneys Office Drug Treatment Alternative to Prison (DTAP) Program. *Justice Research and Policy, 7*(1), 1-25.