CONDUCTING BENEFIT-COST ANALYSES OF JUVENILE COURT JURISDICTION AND OTHER JUVENILE JUSTICE POLICIES

December 21, 2009

Jens Ludwig, Roseanna Ander & Laura Brinkman
University of Chicago

Contact:
Jens Ludwig
University of Chicago
1155 East 60th Street
Chicago, IL 60637
(773) 834-0811
jludwig@uchicago.edu

This paper was made possible by the generous support of the MacArthur Foundation. Jens Ludwig is the McCormick Foundation Professor of Social Service Administration, Law, and Public Policy at the University of Chicago, and Director of the University of Chicago Crime Lab. Roseanna Ander is the Executive Director of the University of Chicago Crime Lab. Laura Brinkman is Research Associate of the University of Chicago Crime Lab. Thanks to Laurie Garduque, Robert Schwartz and Laurence Steinberg for helpful discussions. All opinions and any errors are of course our own.

I. INTRODUCTION

Concern about juvenile crime in the 1980s led most states across the country to enact policy changes that made it easier to transfer juvenile arrestees into the adult criminal justice system (Hahn et al., 2007, p. 2). For example the number of states with statutory waiver laws, which require certain types of crimes to be automatically tried in the (adult) criminal courts, more than doubled from 1979 to 2003, from 14 to 31 (Redding, 2008). Over the past 15 years, fully 13 states have lowered the maximum age of juvenile court jurisdiction to below 18, though this trend has been reversed in some states in recent years. Still, the expansions of transfer laws lead to an increase in youth being tried and sentenced in the adult corrections system. New admissions for youth under 18 into state prisons increased 22% from 1985-2002 (OJJDP, 2006).[1]

Increasingly constrained criminal justice budgets have created a second motivation for transferring juvenile arrestees to the adult system – costs. The average annual cost of detaining a juvenile offender is estimated to be nearly $88,000 per year,[2] nearly four times the estimated annual cost of keeping an adult in prison.[3] It is natural that many policymakers might conclude that shifting young arrestees from the juvenile to adult criminal justice systems could help conserve resources that could be devoted to other pressing social needs.

Yet a key point of our paper is that this type of informal benefit-cost analysis, which focuses on comparing the annual costs per bed in the juvenile versus adult system, is far too narrow. After all, a key *reason* why the cost per detainee in the juvenile justice system is so much higher than in the adult system is because the former is intended to provide more intensive educational and therapeutic services, and also employs higher staff-to-resident ratios to provide a safer, more secure environment for residents and minimize any "brutalization" effects from

[1] http://www.ojjdp.ncjrs.gov/ojstatbb/nr2006/index.html
[2] http://www.justicepolicy.org/images/upload/09_05_REP_CostsOfConfinement_JJ_PS.pdf
[3] http://www.ojp.usdoj.gov/bjs/pub/pdf/spe01.pdf

detention. As a result we might expect, or at least hope, that handling juvenile cases in the juvenile rather than adult systems could potentially lead to lower rates of recidivism, increased schooling attainment, improved health, and greater success in the (legal) labor market. These potential outcomes not only improve the well-being of juvenile arrestees themselves, but also have important consequences for the well-being of the rest of society as well. Focusing only on the difference between the juvenile and adult systems in the costs per bed can lead to misguided public policy decisions that are not in the best interests of the American public.

This "white paper" tries to explain what benefit-cost analysis (BCA) is exactly, why properly-done BCA is so central to the development of juvenile justice policy, why overly narrow or otherwise inappropriately conducted BCA can lead to potentially harmful policies and practices, and what individual jurisdictions can do to help make greater use of BCA. We focus on ways to conduct BCA of different policies that would affect the likelihood that juvenile arrestees are processed in the adult versus criminal justice system, because this has become an issue of particular policy interest in recent years, although almost all of the key points of our paper are relevant for a broad range of juvenile justice policy decisions.

The remainder of our paper is organized as follows. The next section provides a very brief primer on benefit-cost analysis. The potentially great value of BCA is highlighted by work that the Abdul Latif Jameel Poverty Action Lab (J-PAL) at MIT has carried out in a very different context – policy efforts to improve the lives of children growing up in the developing world. Various J-PAL studies have convincingly shown that it is possible to improve schooling outcomes for very poor children through de-worming treatments, hiring extra teachers, and paying families to send their children to school. But simply knowing what programs work is not enough. In a world of scarce resources, the Kenyan minister of education also needs to know

how to allocate funding across competing uses. J-PAL calculations show that each extra child-year of education costs $6,000 from paying families to send their children to school, $60 from hiring extra teachers, and just $3.25 from de-worming treatment (Banerjee and Duflo, 2008). While benefit-cost analysis is often wrongly perceived to have a narrow green-eye-shade focus on saving money, to the detriment of the well-being of children and society at large, the J-PAL example highlights that the goal of a well-done benefit-cost analysis is the *opposite* of that.

The third section of the paper argues that any benefit-cost analysis is only as good as the underlying evidence for the *causal* effects of the policy change that is being considered. It is very difficult to identify this causal effect in the real world, a problem that can lead to unhelpful or even quite harmful policy decisions, as illustrated for example by the well-known case of hormone replacement therapy (HRT). A large body of non-experimental research in epidemiology from the Nurses Health Study found that women who received HRT to deal with the onset of menopause had better health outcomes than women who were observably similar with respect to age and a variety of other characteristics, but who did not receive HRT. On this basis, literally thousands or even millions of women were encouraged to receive HRT by their doctors over the years. Eventually a formal randomized experiment was carried out, which found not only that HRT had no beneficial effects – the trial had to be stopped early because of the adverse health impacts that were found to occur to women from HRT (Hsia et al., 2006, Angrist and Pischke, 2009). Carrying out a BCA of HRT using data from the non-experimental studies of the treatment's impacts on health would have simply reinforced the social harm that resulted from the flawed research claims about HRT's benefits to health outcomes. In this section of the paper we also provide some additional information about various research designs that we believe can be helpful in isolating the policy effects on key outcomes of interest.

The fourth section of the paper discusses the sort of data on both benefits and costs that jurisdictions should collect to be able to translate a strong research design into a successful benefit-cost analysis. Our basic argument is that jurisdictions should follow what is essentially a Willie Sutton strategy, and focus their attention on where the money is. Trying to carry out a truly comprehensive cost-benefit analysis, and collecting detailed data on all aspects of the criminal justice system and subsequent children's outcomes, can easily become an overwhelming task that discourages doing any BCA by local jurisdictions that are already pressed to fulfill their law enforcement and social service obligations. The effort to be comprehensive can lead to paralysis by analysis. We argue instead that agencies should focus their attention on collecting data on those big-ticket items that are likely to be most important for assessing the general net benefits and costs of different juvenile justice policies.

The fifth section discusses the process of assigning dollar values to program impacts on outcomes such as recidivism or schooling attainment that are not already denominated in dollars. We briefly discuss some of the conceptual difficulties of assigning dollar values to intangible outcomes, and then also provide a "price list" drawn from the best available research published in the scientific literature.

The final section concludes with some additional "cook book" advice. Some agencies will have the capacity and interest in carrying out BCA on their own, and so we provide some additional guidance and suggestions for further reading about how to assemble and interpret the evidence that is available about policy benefits and costs. For agencies that would benefit from technical assistance in carrying out BCA, or that would prefer to even outsource the entire enterprise, we provide some information about non-profit organizations that have been

established to carry out this sort of work, including the University of Chicago Crime Lab and the Vera Institute of Justice's Cost-Benefit Analysis Unit.

II. BENEFIT-COST ANALYSIS: A PRIMER

Benefit-cost analysis (BCA) is designed to help policymakers understand which policies generate benefits to society that are large enough to justify a program's costs. In order to compare program benefits to program costs, both sides of the ledger need to be converted to a common metric, which under BCA is dollars. While the basic idea behind BCA is quite straightforward, the application of BCA in practice, including in juvenile justice settings, raises a number of additional complications and potential sources of confusion.

BCA is intended to be an aide to decision making, rather than the decision itself, since policymakers will always – and appropriately – consider other factors besides net benefits as well, including political considerations and the rights of different affected groups (for example Zerbe, 1991).

A good BCA will also enable policymakers to understand the distributional implications of specific policy choices; the usual logic behind BCA is that any policy that generates benefits to society that exceed social costs could in principle enable the "winners" to fully compensate those who "lose" under the policy and still retain some surplus, known in economics as the potential Pareto improvement principle. But as a society we very rarely require this compensation to occur in practice, as witnessed, for example, by the large gains to society that most economists believe occur from international trade, juxtaposed against the wrenching social changes and economic decline of Flint, Michigan and other manufacturing areas in the Rust Belt and across the country. (For an excellent example of a benefit-cost analysis that disaggregates

program benefits and costs to different constituencies, see Long, Mallar and Thornton's 1981 study of the Job Corps program).

While showing the effects of public policy decisions on different societal sub-groups is a necessary condition for any well-done BCA, it is important to recognize that this is not sufficient. A well-done BCA must also still enable policymakers to understand the net consequences of some policy change from the perspective of society as a whole, even if bureaucrats often tend to focus in practice on inappropriately narrow considerations such as the revenue inflows and outflows of their particular agencies (Boardman, Vining and Waters, 1993).

To take an extreme example, consider a policy that would give nuclear facilities the option of dumping radio-active waste into Lake Michigan rather than transport it to Yucca Mountain, Nevada. This policy change would surely produce substantial savings to whatever department within the U.S. Department of Energy is responsible for helping the nuclear industry with disposal of radio-active material. But no one would want policymakers to commit to such a decision without having a full appreciation of the consequences for public health and safety.

A less extreme example occurs when, say, the department of corrections is asked to assess some policy change that requires resource expenditures by that agency in exchange for producing benefits to human service agencies, in addition perhaps to also producing benefits to society as a whole. The solution to this problem presumably depends in part on having higher-level policymakers be the ones carrying out and assessing the results of benefit-cost analyses. Having a governor's office or state legislature or some other entity responsible for all the relevant agencies make the decision about whether the department of corrections should enact whatever policy has been subjected to a benefit-cost analysis is one way to ensure that the

external benefits of this policy change that are generated across different government agencies are "internalized" in the decision making process.

In the case of the juvenile justice system, this means that we need to measure the effects of juvenile justice policies on the behavior and well-being of juveniles and others in society, not just the dollar costs or savings to government budgets. This is particularly important with respect to decisions about whether to handle juvenile arrest cases in the juvenile versus criminal justice systems, since youth behaviors may be affected by placement decisions. As juveniles are generally thought to be less culpable for their crimes, and more amenable to treatment, there is more extensive programming available to youth in juvenile correctional facilities, and a bigger focus on rehabilitation (Scott and Steinberg, 2008). For example one study of educational programming in prisons found that the teacher to student ratio was 1:100 in adult prison (Scott and Steinberg, 2008), compared to the 1:15 ratio recommended by the American Correctional Association for youth involved in the juvenile justice system. Qualitative interviews also suggest that juvenile offenders in the adult prison system generally do not receive any special programming and are treated similarly to adult inmates (citation). Put differently, if the juvenile justice system is working as it should (or at least as we hope), we are trading off higher detention costs in the short term in exchange for improved behavioral outcomes in the future.

Note one implication of this argument – a proper BCA must pay attention to juvenile behavioral outcomes beyond just recidivism. In principle any criminal justice or social policy intervention that effectively reduces someone's propensity to commit crime may also have salutary effects on other behavioral outcomes that represent the flip-side of crime-proneness, such as schooling attainment, mental (and even physical) health, and success in the formal labor market. Ignoring these behavioral impacts on outcome domains besides crime may lead to a

substantial understatement of the benefits of putting youth into more developmentally productive environments. The flip side is that we also need to take seriously the possibility that the longer sentences and harsher detention conditions found in the adult criminal justice system could have some deterrent effect on the criminal behavior of all juveniles in the jurisdiction, which could be mitigated to some degree by policy changes that reduced the likelihood that serious juvenile offenders are transferred to criminal court. We recognize that the question of whether criminals in general, and adolescent offenders in particular, can be deterred by the threat of punishment is controversial. Our own view is that there is a sufficiently strong body of empirical research to at least consider deterrence a logical possibility that must be considered, rather than dismissed outright, recognizing that the empirical basis for deterrence is something less than definitive.[4] Perhaps the best way to empirically account for deterrence effects, which essentially generate spillover effects on other youth, would be to carry out a benefit-cost analysis at the local-area or jurisdiction level, rather than at the individual youth level. We return to this point below.

A related implication is that any proper BCA *must* pay attention to the intangible costs associated with government decisions. Just because the crippling effects of a mental or physical health problem do not show up on a government ledger do not mean that they do not have tremendous negative consequences for the ability of people to take joy out of life, which after all is ultimately the main goal for economic activity overall. Similarly the tangible costs of criminal behavior are often trivial compared to the intangible costs associated with the trauma of injury or having lost a loved one, as we discuss in detail below.

[4] Levitt (2003) reviews empirical studies of the deterrence hypothesis. One particularly relevant data point for present purposes is the finding of Levitt (1998) that the difference in offending rates between juveniles and adults is most pronounced in states where the penalties between the two systems is most pronounced. However Lee and McCrary (2005) use data on exact date of birth and find that there is no detectable "jump" in offending rates at the age of majority. On the other hand ethnographic accounts support the idea that the threat of punishment affects behavior; see for example Cook, Ludwig, Venkatesh and Braga (2007).

Another implication is that we will want to pay careful attention to the long-term as well as short-term costs and benefits of some policy change, because the criminal justice cost differences between the adult and juvenile systems are felt immediately, while any impacts on youth behavior are likely to accrue over time. This raises a complication – no jurisdiction would all else equal be willing to trade a $1 million expenditure today for a payment in 10 years of exactly $1 million, because the jurisdiction could in principle have invested their $1 million today in some interest-bearing bond or other asset and in 10 years have both the original principal and accrued interest to devote to different government activities. Put differently, there is some opportunity cost to government agencies, just as to businesses and consumers, from receiving benefits off into the future. Standard practice in benefit-cost analysis is to discount future benefits or costs to reflect this opportunity cost. Options for carrying out this discounting are described in the final section below for agencies interested in carrying out their own BCA.

We have argued that any proper BCA of policies that affect the propensity of juvenile offenders to be processed by the juvenile versus adult criminal justice system must account for the impacts on juvenile behavior, given that the two systems may differ in the harshness of their detention conditions, the differences in sentence lengths (that could affect general deterrence on the juvenile population at large), and, of course, in the provision of quality educational and therapeutic services to detainees. Of course these features of the adult and juvenile systems vary across jurisdictions. This variation implies that the behavioral consequences (and the system costs) of being processed by the juvenile versus adult system are likely to vary across areas that have different juvenile and adult system characteristics. This possibility means that local policymakers should be very cautious drawing inferences about the benefits and costs of juvenile versus adult processing from national studies or studies drawn from very different types of

jurisdictions, and suggests the potentially significant value of jurisdiction-specific BCA whenever possible.

Another reason that the benefits and costs of processing youth in the juvenile versus adult systems may vary across jurisdictions is variation across areas in the caseload mix, including the average prior criminal history, of juvenile arrestees. For example the psychological literature seems to suggest that cognitive behavioral therapy (CBT) has more pronounced beneficial impacts on those youth who have the most extensive prior history of arrests, though not with a history of violence (Lipsey et al., 2007). This point reinforces our caution about taking "off-the-shelf" BCA analyses from other places and applying them locally, and also suggests the possibility that the benefits and costs of policies that change the likelihood that juvenile offenders are processed in the juvenile versus adult systems may depend on the specific policy mechanisms that we use to achieve that end. For example a policy that lowers the age of majority from 18 to 17 in a jurisdiction will push low-risk as well as high-risk youth into the adult system, while a policy that makes it easier for judges and prosecutors to transfer juvenile offenders into the adult system, holding the age of majority constant, would be expected to at least on average shift a higher-risk caseload of youth into the adult system compared to a change in the overall age of majority. We are not, of course, taking a stand on which of these policies (if any) would be preferable in any way in some normative sense. We only intend to make the point that if there is variation across jurisdictions and specific policies in their benefits and costs to society, then there may be value in trying to carry out benefit-cost analyses that are to the greatest extent possible specific to local jurisdictions and juvenile justice policies.

We close by noting that one closely-related alternative to BCA is cost-effectiveness analysis (CEA), which tries to overcome some of the conceptual and practical difficulties of monetizing

intangible program benefits by simply ranking and comparing programs that reduce crime on the basis of the cost per crime averted. (A proper cost effectiveness analysis would want to at least account for the fact that different types of crime vary in their harm to society, but this is the basic idea). For example a program that required $10,000 in spending for each homicide averted would be more cost effective than an alternative program that required $100,000 in spending per homicide prevented. But this second-best strategy of using cost-effectiveness analysis has the drawback of limiting policymakers to comparing just across policies that try to reduce crime. This is a drawback because it ignores the possibility that even some of the less cost-effective juvenile justice policies could have more favorable benefit-cost ratios than other policies that address problems other than crime, or, conversely, that even the most cost-effective juvenile justice policy changes could have lower benefit-cost ratios than other non-crime policies.

III. MEASURING CAUSAL POLICY EFFECTS

Any benefit-cost analysis (BCA) is only as good as the underlying evaluation evidence that tells us about the causal effects of the policy of interest. Trying to monetize the benefits and costs generated by some policy change may do no good, and may well even do harm, if we are not confident that we understand what the policy does to key outcomes such as criminal justice expenses, as well as youth outcomes like recidivism rates, schooling, and deterrence. At the same time, determining the causal effects of policies on these outcomes is more difficult than is commonly appreciated.

The most common way that people tend to think about causality is to see what happens to some outcome after something happens. This is known in the program evaluation literature as a "before / after" or "pre / post" design, and is what newspaper reporters, policymakers and the public at large do intuitively all the time. For example in New York City in the 1990s under the

Guiliani administration, the NYPD implemented a widely-cited "broken windows" policy that involved stepped-up enforcement of quality-of-life offenses. Crime dropped dramatically in New York during Mayor Guiliani's watch, and most people (including the Mayor) were inclined to attribute the drop to the new policing strategy. The problem with this sort of before / after thinking is obvious once one recognizes that crime dropped almost *everywhere* in America during the 1990s, not just in New York, as shown in Figure 1 (Levitt, 2004).[5]

Another common way that people think about causality is to compare what is going on in different places that have made different policy decisions. Consider, for example, an article by the Toronto *Star* in 2008: "With new laws in Canada that will put more people away, for longer … the cost of jailing is about to go up. If the aim is to reduce crime, Michigan is proof otherwise."[6] The "proof" in this case comes from noting that compared to Minnesota, the state of Michigan has a much higher incarceration rate (around 500 per 100,000, vs. 176 per 100,000) and also a much higher violent crime rate (562 per 100,000, compared to 312 per 100,000); see Figure 2. While this sort of argument is common, it is no more informative than noting that people in the waiting room of doctor's offices are sicker than the general population at large. After all, one reason that Michigan might have a higher incarceration rate is that the state has a more disadvantaged and crime-prone population compared to Minnesota. For example as seen in Figure 2, compared to Minnesota, Michigan has a higher poverty rate (14 versus 10 percent), lower high school graduation rate (83 versus 88 percent), and lower median household income

[5] A more formal analysis of the impact of broken windows policing in New York City comes from comparing crime trends over time in police districts in the city that received more versus less broken windows policing. This sort of analysis suggests that broken windows policing seems unlikely to have contributed much to the city's crime drop (Harcourt and Ludwig, 2006).

[6] Betsy Powell, July 23, 2008, "Has 'mass incarceration' failed?" The Toronto *Star*. http://www.thestar.com/SpecialSections/Crime/article/460767

($48,000 versus $55,000), plus Michigan has a major city (Detroit) with massive amounts of concentrated disadvantage.

Even fairly sophisticated non-experimental studies, that try to statistically control (or "regression-adjust") for observed differences in people, groups or jurisdictions that are being compared can lead to incorrect inferences about what good different policies are accomplishing. The problem arises because even if we control for all observable attributes of, say, Michigan versus Minnesota, there might still be other harder-to-measure characteristics of the two states that affect crime rates.

The case of hormone replacement therapy (HRT) mentioned in the introduction highlights the potential importance of this "omitted variables" problem in practice. Using observational data from the Nurses Health Study, epidemiologists found that observationally equivalent women who received HRT had better health outcomes than those who did not receive HRT. A randomized experiment, which is the "gold standard" for identifying causal effects of some policy or other intervention, found that HRT actually harms women's health on average. The implication is that women in the Nurses Health Study who had voluntarily chosen to receive HRT have unobservable characteristics that have protective effects on their health outcomes, for example perhaps because the women who receive HRT had higher incomes or wealth or health insurance coverage that may have been unmeasured or inadequately measured in the Nurses Health Study and that enabled these women to be able to afford HRT.

One implication is that in order for BCA to serve as a force for good in helping inform juvenile justice decisions, local analysts and policymakers must hold a very high standard of evidence for claims about the causal effects of juvenile justice policies on key outcomes of

interest. Three types of research designs in particular seem promising for juvenile justice applications.

A. Random Assignment

The U.S. Food and Drug Administration (FDA) requires Merck, Pfizer or any other drug company to carry out a randomized clinical trial to determine a drug's benefits before it can be legally sold. A randomized trial isolates the causal effects of the drug on patient health because under random assignment, the active drug treatment group and the placebo control group should on average be identical, with the important exception that the former receives access to the medication that is being studied. Therefore, any differences in average health outcomes across the two groups can be confidently attributed to the causal effects of the drug itself.

Many readers will immediately object – it is unethical to randomly assign juveniles to be processed and punished in the juvenile versus adult systems. And yet that is what local criminal justice agencies do in practice all the time, at least probabilistically speaking. In many jurisdictions, judges and prosecutors have a tremendous amount of discretion in how juvenile arrestees are processed. Put differently, two juveniles with the exact same criminal records can be treated very differently by the criminal justice system simply by virtue of the judge and prosecutor working on their case. And many jurisdictions explicitly assign criminal cases randomly to judges and prosecutors to avoid the potential for corruption.

There is the potential to carry out a "gold standard" evaluation (essentially like a clinical trial in medicine) in any situation where criminal justice actors have discretion in how they handle cases and where cases are randomly assigned to criminal justice decision makers for processing. Suppose, for example, that Judge Charles Hughes waves 30 percent of his juvenile caseload into the adult system, while Judge Harlan Stone waives jut 5 percent of his cases into

the adult system. Suppose also that otherwise, these judges treat similarly situated cases the same way. If cases are indeed randomly assigned to judges, then any difference in the average outcomes of juveniles assigned to Judge Hughes rather than to Judge Stone can be attributed to the fact that the juveniles assigned to Judge Hughes are 25 percentage points more likely to be transferred to the adult criminal justice system. Note that this research design can be employed even if assignment is randomly only conditional on observed case characteristics (for example, if children are assigned to judges serving their local police district, but conditional on police district, assignment of children to judges is random; in this case, the analysis would simply focus on comparing children assigned to judges within the same police district). A similar idea can be applied to study the consequences of *any* juvenile justice practice where actors have discretion and vary in how they use this discretion, so long as cases are randomly assigned to criminal justice decision makers (such as probation officers, etc.) For an example of this sort of study in practice, see Kling (2006).

Another example of naturally-occurring random assignment comes in cases where agencies choose to use random lotteries as a fair way to allocate scarce government services. For example the Chicago Public Schools has a number of unusually desirable magnet, charter or other schools that are in principle open to children all across the city, not just those who live in the school's immediate catchment area. Because the overall quality of the CPS system is quite uneven, it turns out there is substantial excess demand for many of the city's most desirable schools – that is, there are far more children who apply to attend these schools than there are enrollment slots available in the schools. As a fair way to allocate scarce slots in the schools, CPS holds admission lotteries, which has enabled researchers to carry out extremely rigorous studies of the effects of these desirable schools on children's outcomes compared to the other

schools to which children would be assigned if they did not win the admission lottery (Cullen, Jacob and Levitt, 2006). Another implication of this example is that whenever jurisdictions roll out appealing new pilot programs for which there is excess demand, random lotteries could be employed prospectively as a fair way to allocate services, which would have the secondary benefit of enabling jurisdictions to carry out exceptionally rigorous BCA studies of these pilot programs.

B. Regression Discontinuity

A second type of possibility for a good evaluation comes when program services are allocated on the basis of ranking target populations on the basis of some sort of risk or protective indicator or score, and then prioritizing services for populations above or below some given score. The key assumption behind this evaluation approach is that whatever other potential confounding factors might influence the outcomes of interest will vary "smoothly" across the funding or service-assignment threshold. If the only thing that "jumps" discontinuously at the threshold is funding or access to the program being studied, than any "jumps" in the outcomes of interest at the threshold can be attributed to the causal effects of the program itself. A growing body of research in statistics and program evaluation now suggests that this type of "regression discontinuity" design is almost as good as carrying out a formal randomized clinical trial.

An example of this type of research design comes from the paper one of us (Ludwig) published with Douglas Miller to study the effects of the federal Head Start program on poor children (Ludwig and Miller, 2007).When the Office of Economic Opportunity (OEO) was launching the Head Start program in the 1960s, they essentially wound up providing extra Head Start funding resources to the 300 poorest counties in the U.S., which turned out to be all counties with poverty rates in excess of 59 percent. Figure 3 shows Head Start funding rates per

child in the county in 1968 (dollar amounts shown on the y-axis). Counties are ranked on the basis of their 1960 county poverty rate (shown on the x-axis). The vertical line in the graph shows where the county poverty rate threshold was to receive extra Head Start funding (as noted above, equal to 59 percent poor according to the 1960 census). The triangles in each graph represent the mean funding level for groups of counties with similar 1960 poverty rates, the gray bars or "whiskers" around those triangles represent 95% confidence intervals, and the solid and dashed lines that run through the triangles show different parametric and non-parametric regression lines fit to the data. Starting at the left hand side of the graph, we see that as the county poverty rate increases (that is, as we begin to move to the right in the figure), Head Start funding rates increase very slightly, consistent with the idea that counties that are a little bit poor tend to get a bit more Head Start funding just as a matter of course. But then when we reach the vertical line, we see that OEO's decision to provide extra funding to the 300 poorest counties causes a large "jump" or "discontinuity" in Head Start funding levels.

Figure 4 plots government spending on all other social programs, which, like Figure 3**,** ranks all counties according to their 1960 county poverty line. The figure shows that there is no similar "jump" at the Head Start funding level in terms of other types of social spending. The fact that all of the other relevant determinants of children's outcomes vary "smoothly" at the Head Start funding threshold – except, of course, for Head Start funding itself – means that any discontinuous jump in child outcomes that we observe at the Head Start funding threshold can be attributed to Head Start.

Finally, Figure 5 shows that the upward "jump" in Head Start funding levels shown above is mirrored by a downward "jump" in mortality rates to children from causes of death that were screened for by Head Start. Starting with the left hand side of the graph, we see that child

death rates increase as counties become increasingly poorer and poorer. But then when we reach the threshold poverty rate to receive extra Head Start funding (the vertical line) we see a large downward shift in deaths.

A similar type of research design can be applied to understand the effects of being processed in the adult versus juvenile system by, for example, taking advantage of data available on the exact date of birth of arrestees together with knowledge about the local age of majority. For example David Lee and Justin McCrary (2005) use longitudinal felony arrest data from the state of Florida to examine whether the threat of more severe punishments for offenders in the adult vs. juvenile justice system help deter youth criminal behavior. This "deterrence" hypothesis suggests that if we plotted arrests to people based on the number of days before or after their 18th birthday in Florida, we would expect to observe a sharp drop in arrests to people right on their 18th birthday, since youth experience a sharp jump upward in criminal penalties for most offenses at that age. Yet they do not find evidence of such a sharp drop in arrests, which suggests that the extra severity of adult punishments relative to the juvenile system may not have large general deterrent effects on youth.

This sort of research design could also be employed to study the effects of any sort of disposition outcome or aftercare services where assignments are driven by sharp cutoffs in risk and protective scores. For example, imagine that an agency assigns substance use treatment to any youth who has a risk score of 30 or higher. Those youth with scores of 29.9 should be quite similar in practice to those with scores of 30.1 in terms of their overall set of risk and protective factors, with the exception that the latter has access to subsidized substance use treatment assistance. This design may have widespread application because a growing number of juvenile justice around the country are working to limit the amount of discretion about juvenile arrestee

dispositions or after care assignments, using screening assessments and formalized rules to help guide these decisions.

C. Policy Changes and Difference-in-Differences

A third possibility is to take advantage of policy or program changes that affect some groups but not others, which is known in the program evaluation literature as a "difference in differences" study. As Angrist and Pischke (2009, p. 227) note, one of the first studies to use this design was carried out in 1855, which also helps illustrate the simplicity of this design. Difference-in-difference was used to test whether cholera is transmitted by contaminated drinking water rather than contaminated air. In 1849 two of the main water companies operating in London, the Southwark and Vauxhall Company and the Lambeth Company, both drew their water from the polluted Thames River. In 1852, the Lambeth company moved its facilities upriver to reduce the amount of raw sewage taken in with the water. The London districts served by the Lambeth Company showed a sharp decline in cholera deaths compared to the trend observed in parts of the city served by the Southwark and Vauxhall Company (that is, the decline in cholera deaths from cells A to B in Table 1 was larger than the decline in cholera deaths from cells C to D).

Perhaps the biggest challenge to successfully carrying out this difference-in-difference design is to find a control group; that is, a set of people or governmental jurisdiction that does not experience the policy change of interest, and that would have experienced similar trends in outcomes had the treatment group not been subject to the policy change. One way to determine whether one has successfully identified an appropriate control group is to compare trends in outcomes across the two groups before the policy change is enacted (that is, during the "pre-treatment" period). If the "treatment" and "control" groups have different trends in outcomes

even *before* the new program or policy of interest is implemented, then we cannot be sure that any difference in trends after the policy is implemented can be attributed to the policy itself.

The controversy over the federal Brady Act highlights the challenges and potential benefits of implementing this research design. Beginning in 1994 the Brady Handgun Violence Prevention Act required all states to conduct background checks on prospective buyers, thus changing policy for 32 states that previously lacked this requirement. One of us (Ludwig) carried out a difference-in-difference study of Brady's effects, published in the *Journal of the American Medical Association,* that compared mortality trends in the 32 states whose policies were altered due to the Brady Act's requirements ("Brady treatment states") with the 18 states that were exempt from the Brady provisions due to pre-existing sufficiently stringent background check requirements ("Brady control states"). Figure 6 below shows that there were no statistically different homicide trends among people ages 21 and older (Ludwig and Cook, 2000).

Note that the focus on adult mortality in this analysis has two separate motivations. First, people under age 21 were not legally allowed to buy a gun from a licensed federal firearms dealer even before Brady, so the background check might be expected to have much more modest impacts on juveniles compared to adults. Second, juvenile homicide rates had different trajectories in the treatment and control states even before Brady was enacted. As a result, any differences in juvenile homicide trends that we observed between the Brady treatment and control states subsequent to the Brady law in 1994 could not be confidently attributed to the effects of the Brady law itself.

IV. DATA REQUIREMENTS FOR BENEFIT-COST ANALYSIS

We have argued that there is great value in carrying out, to the extent possible, benefit-cost analyses at the local level and for specific juvenile justice policies. Local jurisdictions may

or may not have the interest and capacity to carry out these BCAs themselves, versus outsourcing these analytical activities to other government agencies or non-profits. But at the very least any BCA will require access to administrative and related data that only local jurisdictions can assemble. Even jurisdictions that are not currently actively contemplating BCA would benefit from assembling high-quality data systems to preserve the option value of retrospectively conducting BCA in the future. Note that because many juvenile justice policies affect the likelihood that youth are processed by the juvenile versus adult system, which may differ in both their costs and benefits, proper analysis of juvenile justice policies will require data from both the adult as well as the juvenile systems.

In a world in which collecting and processing data was costless, we would recommend that juvenile and adult systems collect information on anything that might be potentially relevant. But in the real world, government agencies and their community partners are, appropriately, focused on serving their constituents, and so data collection is costly and diverts staff resources from other operational needs. At the same time, data collection may have great benefits by highlighting ways in which agencies can improve the efficiency of their practices. Recognizing these competing short-term versus long-term priorities, we recommend that agencies focus their data-collection efforts on those benefits and costs that are likely to be most important in driving overall net benefit figures.

A. Costs

One important aspect of cost is to reliably assemble and maintain electronic information about the duration of detention for juvenile arrestees, given the sizable cost differences noted above between detention in the juvenile versus adult systems. Because court records usually just report the disposition of the case (what the juvenile is assigned to, not what they actually

experience), it may be necessary to link data from pre-trial detention facilities and the local juvenile and adult detention systems together.

Of course the large difference in average costs between juvenile and adult detention reflects differences in average costs. The set of juveniles who are transferred from the juvenile to adult system (or from the adult to juvenile system) under some given policy change might have costs that are higher or lower than this average, if, for example, those youth whose status is affected by the policy change are unusually likely to receive services. For this reason it will be important for jurisdictions to also try to assemble electronic records on whatever special services that youth receive within the detention facility, focusing on the big-ticket items related to the most costly services such as schooling, substance or mental health treatment, and physical health treatment.

A more subtle issue is that the marginal costs (or cost savings) from changing the number of youth in a particular setting may be much lower than the average cost if there are important fixed costs associated with the setting (for example, if a large share of the cost of maintaining core staffing and physical upkeep of a juvenile justice facility is independent of the number of youth). It will be important to consider whether a given policy causes a large enough change in the number of youth in a given setting to enable policymakers to close facilities or parts of facilities to realize average cost savings, or whether the changes in juvenile placements are so small that we should think about marginal cost savings instead.

The other key issue we consider most relevant on the cost side of the equation has to do with collecting data on the duration and intensity of probation supervision, and services that probation may assign, supervise or facilitate, that juvenile offenders receive in the juvenile versus adult settings.

Based on our initial examination of the research literature, including previous attempts to carry out benefit-cost analysis of adult vs. juvenile system processing, we expect most of the other cost differences (such as the differences in juvenile vs. adult court system processing itself) to be of secondary importance from the perspective of carrying out a benefit-cost analysis, that is, unlikely to generate much impact on the overall assessment of a given policy change's net benefits or costs to society. [need to confirm]

Finally, agencies should remember that just because they do not receive a separate bill from a vendor for some specific input does not mean that this input does not have a cost. For example having 100 hours of extra staff time devoted to trying to ensure youth can enroll back in the public school system is a cost, even if the personnel making those calls and school visits were already employed by the agency, because of the opportunity cost of what other productive activities those staff-hours could have been devoted to instead.

B. Benefits

One of the most important potential benefits of juvenile justice policies is to reduce recidivism rates among juvenile arrestees. Changes in the frequency of criminal behavior will surely be a major contributor to any benefit-cost analysis in this area, given that the total social costs of crime to the United States is estimated to be on the order of $1 to $2 trillion per year (Anderson, 1999, Ludwig, 2006). In the famous Perry Preschool early childhood educational program, fully 70 percent of the dollar value of the program's benefits on participants came from reductions in criminal behavior (Belfield et al., 2006).

While most "criminal careers" are relatively short, a subset of youth who are involved in crime during adolescence persist in their criminal ways into adulthood. To the extent to which policy decisions affect the persistence of these criminal careers, it will be important for local

jurisdictions interested in carrying out BCA of their juvenile justice policies to access adult as well as juvenile arrest records (or facilitate access to such records by outside researchers).

A different challenge is that privacy protections are typically much greater for juvenile than for adult criminal justice records. For juvenile justice agencies that are carrying out their own BCA, this should not pose a problem, but for agencies that wish to outsource their BCA activities, securing permission to share juvenile arrest data with outside research teams can be complicated. One possibility is to take advantage of the fact that for analytic purposes, researchers do not require any individual identifying information like name, exact date of birth (DOB) or social security numbers (SSNs) once relevant administrative data have been linked together. Analysis of different public policies in the city of Chicago, for example, are greatly facilitated by the fact that Chapin Hall at the University of Chicago has permission from relevant government agencies to access all the relevant administrative databases (criminal justice and social service records), and subject to government agency approval, can link together data systems and provide researchers then with de-identified linked administrative datasets.

As argued above, schooling outcomes may also be affected by many juvenile justice policy decisions. Many, but not all, public school systems around the country maintain electronic, longitudinal student-level records that include data on enrollment status, absences, grade retention, grades, achievement test scores, graduation status, and in some cases disciplinary actions as well. For obvious reasons access to these school records is closely guarded under federal regulations, specifically the Family Education Rights and Privacy Act (FERPA). FERPA laws require schools to have written permission from parents or eligible students in order to release any information about the student's record, although certain organizations conducting studies on behalf of the school can be granted access to administrative

student records for research purposes, and FERPA data access rules are somewhat less strict regarding use of student records that are de-identified (that is, have names and other specific individual identifying information removed).

A third key data source for measuring the benefits of different juvenile justice policies are state earnings records, maintained by state unemployment insurance (UI) systems.

A final outcome domain that could potentially figure importantly in the net benefits and costs of different juvenile justice policy decisions, but could be difficult to access, are health records. Individual patient records from doctor offices or hospitals are closely guarded by HIPAA, which provides federal protections for personal health information held by covered entities. An alternative approach is to take advantage of state or federal data systems for medical services received by people on Medicaid, which, given the strong inverse correlation between income and propensity to be involved in crime, may capture health services utilization and health outcomes for a disproportionately large share of those youth who would be affected by juvenile justice policy changes. At the very least local juvenile justice agencies might be able to measure impacts on mortality outcomes by matching information on youth to either local death certification information or to the census of death certificates maintained by the National Center for Health Statistics National Death Index.

Agencies might have different options for measuring mental health and related developmental outcomes aside from medical records. For example under the Individuals with Disabilities Education Act (IDEA) schools are required to create Individualized Education Plans (IEPs) for students with developmental disabilities and other learning challenges, that could provide one source of information about policy impacts on this domain. Of course carrying out original surveys would be far better still for this purpose, but also far more expensive.

V. ASSIGNING DOLLAR VALUES TO BENEFITS AND COSTS

Usually the main costs of juvenile detention as noted above – detention, probation supervision, and whatever services are received in either setting – are already denominated in dollar terms, and so there is no complicated process required to try to monetize non-monetary outcomes. If local agencies cannot pull together dollar values for all of the relevant costs in their own jurisdiction, they can always draw on cost estimates from other jurisdictions as a second-best approximation recognizing that there could be some differences in costs across areas.

The benefits of any given policy can be much more difficult to quantify than the costs, especially given that aside from earnings, almost all of the key benefits to different juvenile justice policy decisions arise in youth behaviors that are measured in non-dollar terms. From our previous conversations with different juvenile justice practitioners and academics, we expect there to be considerable controversy about whether it makes sense to include the value of the intangible costs of crime in a benefit-cost analysis. We believe very strongly that it would be a grave mistake to exclude the intangible costs of crime in a benefit-cost analysis. Even though these intangible costs are difficult to monetize (as discussed below), there is no question that they account for the most important share of the overall social cost of crime. It is true that some policymakers may currently be inclined to put less weight on intangible than tangible costs of crime. But in our view this is an argument for the need for an intensive educational campaign, rather than an argument to carry out incomplete benefit-cost analyses that may ultimately lead to policy decisions that make society as a whole worse off.

Note also that if local policymakers remain resistant to considering the full social costs of crime, an alternative possibility is to carry out cost-effectiveness rather than benefit-cost analysis, and rank policies that seek to reduce crime in terms of the cost per crime averted.

A. Crime

Two well-known strategies exist that try to quantify the dollar value of the benefits from reductions in crime: the "bottom up" or "ex post" perspective, which focuses on itemizing and valuing the "cost" to society of crimes that have already occurred to identifiable victims; and a "top down" approach that attempts to generate a single number that reflects what the public is willing to pay (WTP) for a decline in the risk of crime victimization in the future. The latter is far more appropriate for benefit-cost analysis than the former.

To understand why one approach is conceptually preferable to the other, let us first consider what types of policy decisions are intended to be guided by benefit-cost analysis. Some local decision maker is trying to decide whether to incrementally change the likelihood that juvenile offenders are processed in the juvenile versus adult system (or to change levels of after care provision, or access to mental health treatment in the juvenile justice system, etc.) The main cost to society comes from additional expenditures today on some change in service mix. From the perspective of a local policymaker, these resources could have gone instead to other pressing uses such as schools, roads, public transportation, snow removal, garbage collection, homeless shelters, etc. The public good that citizens receive in exchange for devoting extra resources to some policy change designed to help juvenile offenders is some reduction in the risk that they or people they care about will be victimized by these youth in the future. In order to compare the value of this benefit to the costs, we need to convert these benefits to dollar terms, and the appropriate way to do that is to measure the sum of what people in the community are willing to pay (WTP) for changes in crime victimization risk.

The problem with the "bottom up" or "ex post" perspective is that it either doesn't make any sense, is not useful for policy purposes, or both. The intangible values of the loss to "quality

of life" is calculated in the above table by using the statistical value of life and subtracting how much life was reduced due to a particular crime, and then adjusting this estimate to the age of a typical victim (Cohen, 2005). The ex post perspective also tries to answer the problem of how to value intangible costs by relying on jury decisions, but that just pushes the conceptual problem back a step: How do juries derive their dollar values for these costs? One possibility would be to try to identify the dollar amounts required to make victims whole, or what economists call "willingness to accept" (WTA). But anyone who has lost a parent, child or spouse to crime would say that no amount of money would ever compensate for their loss, which for BCA purposes would in turn imply that we should be devoting every dollar of GDP to crime prevention (since the benefits measured in this way would be infinite). When one of us (Ludwig) teaches BCA in his crime policy class at the University of Chicago Law School and ask how juries come up with victim payments to compensate for intangible crime costs, most law students answer "the juries just make it up," which we suspect comes close to the truth.

The more conceptually appropriate way to think about the costs of crime is what the present paper in this symposium calls the "top down" approach, but which one of us (Ludwig) together with Duke professor Philip Cook prefer to term the "ex ante" perspective. The ex ante perspective defines the social costs of crime as what the public in the aggregate is willing to pay (WTP) to reduce the amount of crime that would arise sometime in the future, and so nicely corresponds to the resource allocation problem facing policymakers

But even after we have settled on the ex ante perspective as the conceptually appropriate way to define what we mean by the costs of crime, there remain a number of difficult measurement challenges. A number of studies have tried to estimate WTP for changes in crime risks by looking at data from housing markets, and specifically looking at what people are

willing to pay for houses in safer neighborhoods in what are known as “hedonic home price” regressions (Rosen, 1988; Cook and Ludwig, 2000; Cook and Ludwig 2001; Cohen et al, 2001; Cohen et al, 2004). But isolating the effects on house prices of safety versus other hard-to-measure home and neighborhood attributes is extremely difficult in practice, since safer neighborhoods tend to have better amenities along a wide variety of other dimensions as well, not all of them easily measured in available public-use data systems, and because home quality may also be somewhat better in safer neighborhoods as well (if only because people in safer neighborhoods may have more money and incentives to try to keep up their homes). Moreover, what someone is willing to pay to live in a 10% safer location understates what they would be willing to pay for a new police program that reduced crime citywide by 10%, since people will put some value on the improved safety of other city residents as well. So estimates for the safety / price gradient in the housing market will likely understate societal WTP for crime control, even if we were not concerned about the possibility of omitted variable bias in home price regressions.

The most common alternative to looking at actual housing market data is to use survey methods to ask people to respond to hypothetical market scenarios, known as the contingent valuation (CV) approach. But this assumes that people have well-formed preferences for safety and are capable of thinking about marginal changes in crime victimization risks. It is possible that these assumptions are met, since most people do have some first-hand experience thinking in at least a general way about crime probabilities in deciding where to buy or rent a place to live, but at the end of the day who really knows. Environmental economists have developed a large literature trying to learn more about whether CV “works” in that application, by for example seeing how WTP responses vary by how the questions are phrased or sequenced or preceded by the provision of different amounts or types of background information, and by trying to construct

scenarios in which WTP survey responses can be benchmarked against actual behavior (Hanemann, 1994; Portney, 1994; Li, Lofgren, and Hanemann, 2002; Carson et al. 2003; Wang and Whittington, 2004). As far as we know there is no similar research program underway in the area of crime, even though it would in our view have tremendous social value.

Table 2 below presents estimates for the costs of different types of crimes using both of these approaches, taken from Miller, Cohen and Wiersema (1996) and Cohen (2005). Despite the conceptual limitations of the ex post perspective, this has been standard practice in the economics of crime literature. Only recently have the more conceptually appropriate ex ante cost of crime estimates become available, and unfortunately these types of ex ante cost-of-crime estimates are only available at present for just a subset of crimes. We present both types of cost estimates in Table 2 for whatever crimes each type of estimate is available.

Two things stand out from Table 2. First, estimates for the costs of crime derived from the ex ante approach (using contingent valuation surveys to ask people what they would pay for a change in the risk of crime victimization) are substantially higher than those derived from the ex post perspective of adding together the different identifiable after-the-fact costs associated with crimes and then using jury award information to value the intangible costs of crime. For example, the cost per homicide under the ex post perspective is estimated to be (in 2009 dollars) around $4 million, while the cost estimate using the ex ante approach is over $10 million, or 2.5 times as high. The disparity is even more dramatic for robbery, for which the ex post perspective yields an estimate of only $11,000, while the ex ante perspective yields an estimate more like $250,000. The robbery example makes clear the limitations of the ex post perspective, since anyone who has ever been held up at gunpoint (or knows someone who has been) will appreciate the substantial intangible consequences of that experience. A second key feature of Table 2 is the

dramatic difference in the social costs of crime for homicide relative to any of the other types of crime. One implication is that even small program impacts on homicide rates can have very disproportionately large impacts on the total net benefits of the program, and so benefit-cost analyses of programs with crime impacts may be quite sensitive to how one monetizes the social costs of homicide in particular.

We encourage juvenile justice agencies carrying out benefit-cost analyses of different juvenile justice policies to carry out their calculations in various ways to determine how sensitive their conclusions are to the choice of cost of crime estimates. One obvious sensitivity analysis is to carry out the BCA first using the ex post cost estimates from Table 2, and then instead using the ex ante cost estimates for those crimes for which these alternative estimates are available. A different sort of sensitivity analysis that is worth carrying out is to examine the sensitivity of the estimates to how one monetizes homicide specifically, for example by examining what happens when the social cost of homicide is assumed to be, say, just twice the value of the next-most-costly crime, rape (see Kling, Ludwig and Katz, 2005).

B. Schooling

Table 3 below presents estimates for the benefits to individuals and society as a whole from policy efforts that successfully increase schooling attainment. The first panel shows the dollar value from increasing high school graduation rates. For example the total social benefit from preventing a white male from dropping out of school is $203,000 in lifetime present value increased earnings and $27,900 from reductions in criminal behavior. In principle we would also want to include the total dollar value of the benefits in health improvements associated with increased schooling attainment, but Levin et al. (2007) from which these numbers are drawn only report the present value of the declines in spending by government health programs (Medicare

and Medicaid). From the perspective of society as a whole reductions in social program participation just represent changes in transfers across parts of society, rather than real net benefits or costs, but from the narrower perspective of taxpayers, increases in high school graduation reduce lifetime present value spending on social programs by $1,200 and $3,300 for white and black males, respectively, and $5,000 and $9,000 for white and black females. Note that we are fairly confident that these reflect the actual causal effects of increased schooling attainment on these different benefits because these estimates are derived from a number of studies that use creative "natural experiments" to isolate this causal relationship, for example by taking advantage of changes over time in state laws regarding compulsory schooling requirements (for example see Card, 1999).

Note also that there is one subtle issue here regarding double-counting of program benefits. Imagine that we have juvenile justice policy change that reduces crime without changing high school graduation rates. The cost of crime figures from Table 2 would then be the appropriate numbers to use. But suppose that we had a policy change that both increased graduation rates and reduced crime. We could not be sure how much of the change in crime was due to the direct effect of the policy change on criminal behavior, versus how much of the change in crime was due to the fact that the program increased schooling attainment and that more schooling causes people to commit fewer crimes. In such cases, we recommend that analysts take the conservative approach of avoiding double-counting problems by using Table 2 to monetize the value to society from reductions in crime, and then subtracting out the benefits of increased schooling attainment due to less crime shown in the top panel of Table 3.

D. Earnings

Some juvenile justice policy changes could potentially affect lifetime earnings prospects through some mechanism other than by increasing schooling attainment. For that purpose, Table 4 presents the present value of lifetime earnings going back to age 18 for different population sub-groups by race and age, taken from Levin (2007). Agencies that have collected earnings data for youth during the first year after their interaction with the criminal justice system and find, for example, that some juvenile justice policy change increases earnings by 5% that first year could use the average annual earnings figure reported in Table 4 to calculate just the first-year gains, and under a sensitivity analysis assume that the earnings gains persist (i.e. that lifetime earnings increase by 5%) and use the present value of lifetime earnings figures.

E. Health

In principle some juvenile justice policy changes could also have some impact on health outcomes that are not directly mediated by policy effects on schooling, in which case agencies carrying out their own benefit-cost analyses would want to add in the dollar value of the health improvements. As noted above, agencies are likely to have better luck obtaining administrative data on mortality (death) than on morbidity (health status). The crime table above noted that the value per life from homicide is around $10 million, which is a reasonable number to use for declines in mortality more generally. Changes in morbidity are more complicated to value. Health economists have tried to measure the dollar value of a given year of life in perfect health, known as a Quality Adjusted Life Year (QALY), with estimates from the literature that range from $50,000 to $100,000 in social value per QALY. One complication for local agencies trying to account for health impacts is that even if they are able to obtain Medicaid records to look at impacts of juvenile justice policy changes on morbidity, the Medicaid data will report outcomes in terms of increased or decreased doctor or hospital visits for specific health problems, which

would then need to be converted into QALY terms (defined on a scale from 0 to 1 as a share of perfect health), and then monetized. This is probably beyond the scope of what local agencies would have the analytic capacity to carry out on their own.

F. Valuing Policies Rather than Outcomes

An alternative approach for trying to assess the dollar value to society from choosing one juvenile justice policy versus another is to ask the public what they would pay to support each policy option, rather than trying to assess the effects of choosing one policy over another on outcomes such as crime, schooling and health and then monetizing those outcomes. Excellent examples of this approach are presented in Nagin et al. (2006) and Piquero and Steinberg (in press). These studies ask respondents whether they would be willing to pay a specified amount per year in extra taxes to support a program that would provide rehabilitation programming for juveniles arrested and sentenced to a year in detention for robbery, noting that such programs reduce youth crime by 30% and may also increase high school graduation and employment rates (but the magnitude of impacts on other outcomes besides recidivism are not specified). Respondents are also asked whether they would be willing to pay a specified amount of additional taxes in order to double the sentence length (from one to two years) for juveniles arrested and sentenced for robbery.

The upside of this approach is that in principle it allows the public to reflect its valuation of different means, not just ends, although Steinberg and Piquero (in press) note that public opinions about different juvenile justice policy options can be sensitive to question wording. One potential drawback of this approach is that it assumes that survey respondents have some understanding about the effects of choosing one policy versus another on the full range of outcomes that might be affected beyond recidivism, such as schooling and employment. But

perhaps a more basic limitation for present purposes is that most agencies are probably unlikely to be able to carry out original contingent valuation surveys of their own.

VI. CONCLUSION: A HOW-TO GUIDE

Perhaps our most important bottom line recommendation to local agencies is: First, do no harm. We have argued that implicit or explicit benefit-cost analyses that focus too narrowly on only a subset of relevant benefits and costs (such as only the monetary consequences to government budgets) can lead to misguided policy decisions that could in the end make society worse off. We have also argued that benefit-cost analyses that are built around weak evidence about what different policy or programmatic alternatives actually accomplish (hormone replacement therapy for women being a canonical example in the health area) can also lead to policy decisions that harm social welfare.

What should agencies interested in supporting a socially productive benefit-cost analysis actually do? One possibility is to ask their research and data departments to carry out an original benefit-cost analysis, perhaps drawing in part on existing estimates in the research literature. A second possibility is to outsource. In what follows we provide practical guidance about both options.

A. Carrying out benefit-cost analysis

Entire books can (and have) been written about how to carry out benefit-cost analysis, such as the excellent text by Anthony Boardman, David Greenberg, Aidan Vining and David Weimer (2006).[7] Any short treatment would not do justice to the various complications that arise in carrying out benefit-cost analysis, so agencies that will carry out benefit-cost analyses on their own are encouraged to supplement our high-level, step-by-step guide here by consulting more in-depth treatments of the details of benefit-cost analysis as well.

[7] Other good but slightly older books on benefit-cost analysis include Thompson (1980) and Gramlich (1981).

Some of the key questions to consider, in order:

1. What is the policy question? It doesn't make sense to ask "what is the benefit-cost ratio of handling juvenile arrestees in the adult versus juvenile systems," because no one is considering outright elimination of the juvenile justice system. In the real world policy change is almost always incremental, and would at most change the proportion and perhaps mix of juvenile arrestees who are handled in one system versus the other. Since the effects of adult vs. juvenile processing may vary across youth (and jurisdictions), what is most relevant for public policy is to understand the benefits vs. costs of making some concrete, well defined change in juvenile justice policy.
2. What is the research design? A good program evaluation is the foundation for any good benefit-cost analysis. We have identified in this paper several research designs that have the potential to generate rigorous empirical evidence about program impacts, several of which take advantage of criminal justice system practices that are common in many jurisdictions. Applying benefit-cost analysis to estimates that have not adequately identified the causal effects of some juvenile justice policy or program of interest may wind up doing more harm than good.
3. What if I don't have a good research design? Then agencies should use off-the-shelf estimates for the causal effect of their policy of interest, ideally from a study that carefully implements well one of the research designs we describe above. Using off-the-shelf estimates for program impacts is not a costless solution to the program evaluation problem, given the possibility that program impacts may vary across areas. For example a policy that increases the likelihood that juvenile arrestees are waived from the juvenile to adult systems may have relatively less pronounced behavioral consequences for youth in

areas where the difference between the adult and juvenile systems in terms of the services they offer and their general overall living conditions are more similar.

4. What are the policy's costs? We suggest agencies focus on the big ticket items that are likely to drive program costs, such as trial costs, detention (and related services), and after-care services. Remember also to account for inflation when assembling program costs from various years, so that everything is reported in constant dollars from some common year.[8]

5. How to measure program impacts? We have argued that one of the key consequences of juvenile versus adult processing has to do with the implications for the long-term life chances of youth – their schooling and employment outcomes, as well as their risk of recidivism. Because surveys are expensive, criminal justice agencies should focus on matching their arrest and court records with student-level school records maintained by their local public school systems as well as earnings records maintained by each state's department of labor as part of the unemployment insurance (UI) system, and social program participation rates (from the perspective of society as a whole this is just a transfer from one member of society to another, but changes in participation rates can have distributional consequences). Because we care about long-term program impacts, these administrative data should be obtained for as many years as possible to track people's outcomes for as long as possible.[9]

[8] Fortunately the U.S. Department of Labor's Bureau of Labor Statistics has an on-line tool that makes these inflation adjustments extremely easy. (http://www.bls.gov/data/inflation_calculator.htm) For example a program that cost $1,000 per child in 1995 would, after accounting for inflation, cost $1,418 today (2009 dollars).

[9] This step will require at the very least linking together administrative data from the relevant court systems, detention facilities, probation, public schools, whatever law enforcement agencies capture juvenile and adult records, and even quarterly employment and earnings information from state unemployment insurance (UI) systems. These data are all highly confidential, but fortunately no individual identifying information is required for analysis once the data systems are linked. Because computing memory and processing speeds are now so good, even with desktop PCs, handling very large data files should not be a problem for most agencies that have some capacity to do

6. How do I monetize these program impacts? Use the price list provided in the previous section. Be careful to avoid double-counting. For example imagine a program that was explicitly designed to improve lifetime labor market earnings for juvenile arrestees by helping them to get a high school diploma, and suppose that we have rigorous estimates for the causal effects of the program on both graduation rates and earnings. Local agencies should note that our figures for the monetized value of the benefits of getting a degree already incorporate the gains in lifetime earnings, and so should not also add in again the dollar value of the program's direct impacts on earnings.
7. Do I have to worry about the fact that most of my program costs are incurred up front, while most of the benefits from improved youth behavior accrue over time? Yes. A dollar today is worth more than a dollar received a year from now because of the opportunity cost of foregoing the chance to have invested that dollar and accrued interest earnings. So all program benefits and costs should be converted to their present value using a discount rate of 3 or 4 percent (citation here).[10]
8. What do I do with these benefit-cost numbers once I have them? As we argued in the introduction, a good benefit-cost analysis enables the reader to understand the distributional implications of a proposed policy change, not just whether the net benefits

---

analysis on their own. Agencies doing their own analyses might aspire to assembling administrative data that go back as far as is relevant in order to be able to examine both how policy effects change over time (as for example programs mature and are refined by administrators), and how policy effects on individual youth change as time passes from when the youth were in the system.

[10] The basic idea is that $1 invested today would generate $1.03 a year from now at a discount rate of 3%, or put differently, the present value of receiving a benefit of $1.03 a year from now at a 3% discount rate is [$1.03 / (1.03)] = $1. The present value of receiving a benefit of $1 one year from now is [$1 / (1.03)] = $0.97. The present value of receiving $1 two years from now is [$1 / (1.03)] * (1 / 1.03) = [$1 / $(1.03)^2$] = $0.94. The present value of receiving $1 a total of N years from now is [$1 / $(1.03)^N$]. Microsoft Office Excel has a function which easily calculates the Present Value of future benefits. The items you will need to know in order to use this function to compute the Present Value of future benefits are the interest rate/discount rate, the length of time you will receive savings from the program or policy, and how often those savings will accrue over the time period. For example if your savings are estimated to be $10,000 every year for three years with an annual rate of 3% the PV function to put into Excel would be: =PV(3%, 3,10,000). The calculated amount, $28,286.11, is the present value of your future stream of benefits for the program.

to society exceed costs. The results of benefit-cost analysis should be presented in a table that has rows for different benefit and cost line items (in present-value terms), and columns that show how these benefits and costs accrue from the perspectives of (a) juvenile arrestees (b) what economists usually call "taxpayers" but by which they mean everyone else besides the juvenile arrestees and (c) society as a whole. (For good examples see Long, Mallar and Thonrton, 1981, or Gramlich, 1981, p. 172). Overall benefit-cost ratios should be presented for each of these perspectives as well.

9. It seems like our benefit-cost analysis requires a lot of assumptions; do we need to carry out some sensitivity analyses to see if our assumptions matter? Yes. For many of the judgment calls required in carrying out benefit-cost analysis there may be no clear correct answer, so the best that can be done is to see whether the conclusions about a program's net impacts are sensitive to one's decisions. Some of the most important assumptions to worry about include:
    a. How one monetizes the social costs associated with criminal behavior, since these figures can be very sensitive to how we monetize the most rare, but socially costly, forms of criminal behavior, and in particular homicide (For an example, see Kling, Ludwig and Katz, 2005).
    b. Sometimes data will not be available for as long a time period as the agency suspects the program might have impacts for. In this case it is useful to carry out sensitivity analyses to see whether the benefit-cost picture for a policy change changes substantially when one assumes that program impacts do or do not persist over time, and under different assumptions about the rate at which program impacts decay over time off into the future.

c. Sometimes the program benefits or costs can be different for small-scale versus large-scale changes in policy. For example many detention centers involve large fixed costs to operate that do not change very much with incremental changes in the detention population. So the cost savings to the government from a policy that would slightly scale back the detention population would be modest, equal to just the marginal cost per detainee, while larger-scale policy changes that allow states to start closing down facilities would generate larger benefits per reduced detainee, since the government realizes average cost savings. On the other side, many social service interventions tend to have larger impacts when delivered at small scale with high quality and fidelity, and tend to generate smaller impacts when delivered at a larger scale.

d. The question of whether to monetize benefits where the estimated program impacts are not statistically significant. Scientific journals usually require program estimates to be statistically significant at the 95% level in order to be considered real impacts rather than statistical flukes, and tend to treat impacts that do not meet this statistical significance threshold as a zero. The reason for that practice is to reduce the risk of "false positive" statistical results, but one cost of that practice is to tolerate a large number of "false negatives" – that is, to treat as zeros some true program impacts that just cannot be estimated with enough statistical precision. The 95% threshold for scientific journals has a strong built-in bias towards the status quo, and so it is useful for benefit-cost analyses to show what happens when impacts that are not statistically significant are also considered.

B. Outsourcing

An alternative solution strategy for local government agencies interested in benefit-cost analysis is to call for help. There is a growing movement within the crime research community to make greater use of rigorous experimental or quasi-experimental research designs to evaluate the causal effects of different criminal justice changes, including changes to juvenile justice system practices. Examples of research centers committed to this sort of activity include, but are not limited to, the University of Chicago Crime Lab (crimelab.uchicago.edu), the Jerry Lee Center of Criminology at the University of Pennsylvania (www.sas.upenn.edu/jerrylee), and the University of California at Berkeley Center for Criminal Justice (www.law.berkeley.edu/bccj.htm), all of which include as a core part of their mission the goal of collaborating with policymakers and practitioners to advance our understanding of various options for preventing crime. In addition, the Vera Institute in New York City has received funding from the U.S. Department of Justice to form a Cost-Benefit Analysis Unit (www.vera.org/centers/cba) to provide technical assistance to local jurisdictions about how to assign monetary values to key program impacts and implement benefit-cost analysis more generally.

Whatever approach is used, we do sincerely hope that more juvenile justice and other criminal justice agencies around the U.S. will begin to make greater use of rigorous program evaluation and benefit-cost analysis as part of their policy decision-making process. America currently has a dramatically higher incarceration rate than any other nation in the world, and yet our overall rates of violent and property crime (excepting homicide) is not that different from what we see in most other developed nations. There may well be the possibility of using more systematic data-driven approaches to reduce our crime rate while simultaneously reducing the human and social costs that result from the government's efforts to control crime.

REFERENCES

Angrist, Joshua D. and Jorn-Steffen Pischke (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Banerjee, Abhijit V. and Esther Duflo (2008) "The experimental approach to development economics." Cambridge, MA: National Bureau of Economic Research Working Paper 14467.

Boardman, Anthony, Aidan Vining, and W.G. Waters (1993) "Costs and benefits through bureaucratic lenses: Example of a highway project." *Journal of Policy Analysis and Management*. 12(3): 532-555.

Cook, Philip J. and Jens Ludwig (2006) Aiming for Evidence-Based Gun Policy. Journal of Policy Analysis and Management. 25(3): 691-736.

Cullen, Julie Berry, Brian A. Jacob, and Steven Levitt (2006) "The effect of school choice on student outcomes: Evidence from randomized lotteries." *Econometrica*. 74(5): 1191-1230.

Gramlich, Edward M. (1981) Benefit-Cost Analysis of Government Programs. Prentice Hall: Englewood Cliffs, NJ.

Hahn, Robert et al. (2007) "Effects on violence of laws and policies facilitating the transfer of youth from the juvenile to the adult justice system: A report on recommendations of the task force on community preventive services." *Morbidity and Mortality Weekly Reports*. November 30, 2007. 56(RR-9).

Hsia, Judith et al. (2006) "Conjugated equine estrogens and coronary heart disease: The Women's Health Initiative." *Archives of Internal Medicine*. 166: 357-365.

Kling, Jeffrey R. (2007) "Incarceration length, employment and earnings." *American Economic Review*. 96(3): 863-876.

Long, David A., Charles D. Mallar, and Craig V.D. Thornton (1981) "Evaluating the benefits and costs of the Job Corps." *Journal of Policy Analysis and Management*. 1(1): 55-76.

Lott, John (2000) *More Guns, Less Crime (2nd Edition)*. Chicago: University of Chicago Press.

Ludwig, Jens and Philip J. Cook (2000) "Homicide and Suicide Rates Associated with Implementation of the Brady Handgun Violence Prevention Act." *Journal of the American Medical Association*. 284(5): 585-591.

Ludwig, Jens and Douglas L. Miller (2007) "Does Head Start improve children's life chances? Evidence from a regression discontinuity approach." *Quarterly Journal of Economics*. 122(1): 159-208.

Nagin, Daniel S., Alex R. Piquero, Elizabeth S. Scott, and Laurence Steinberg (2006) “Public preferences for rehabilitation versus incarceration of juvenile offenders: Evidence from a contingent valuation survey.” *Crime and Public Policy*. 5(4): 301-326.

Piquero, Alex R. and Laurence Steinberg (forthcoming) “Public preferences for rehabilitation versus incarceration of juvenile offenders.” *Journal of Criminal Justice*.

Steinberg, Laurence and Alex R. Piquero (forthcoming) “Manipulating public opinion about trying juveniles as adults: An experimental study.” *Crime and Delinquency*.

Thompson, Mark S. (1980) Benefit-Cost Analysis for Program Evaluation. Sage Publications: Beverly Hills, CA.

Zerbe, Richard O. (1991) “Does benefit cost analysis stand alone? Rights and standing.” *Journal of Policy Analysis and Management*. 10(1): 96-105.

Figure 1
Homicide trends in New York City and other selected cities, 1990-2000

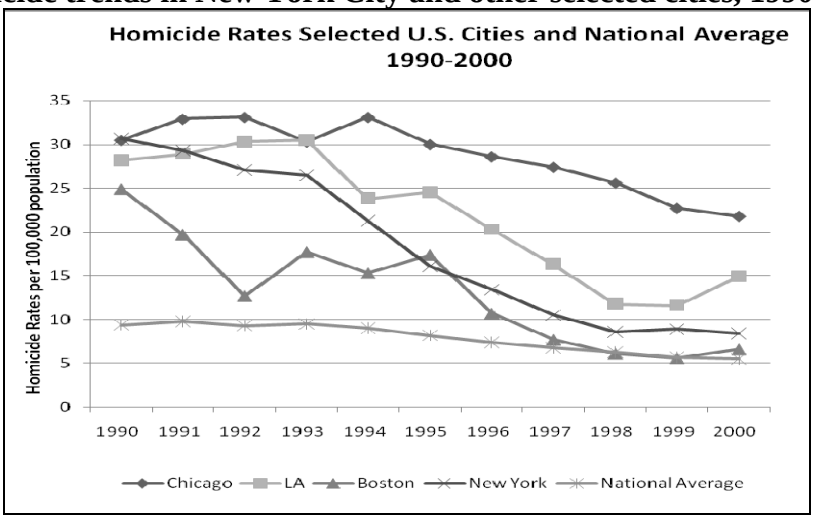


Figure 2
Comparison of Michigan and Minnesota Incarceration Rates, Crime Rates, and Socio-Demographic Characteristics

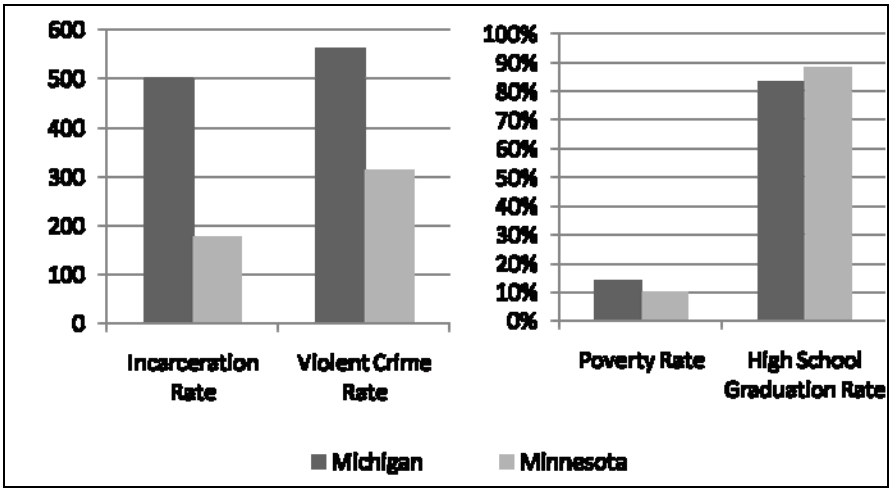
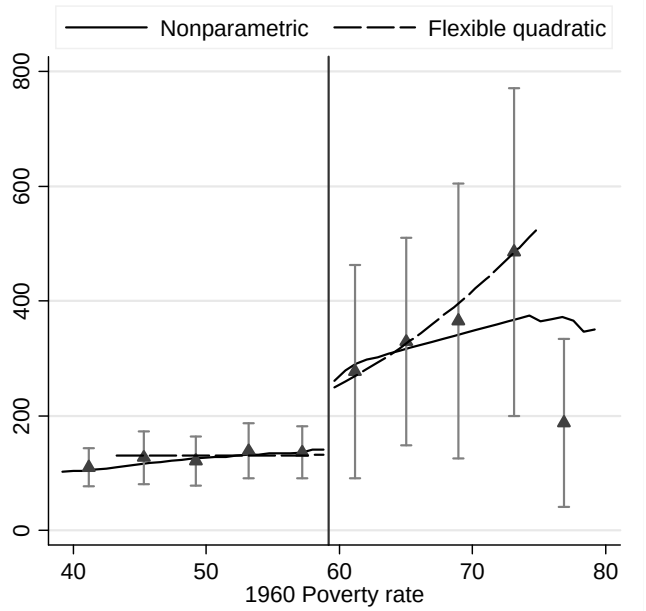
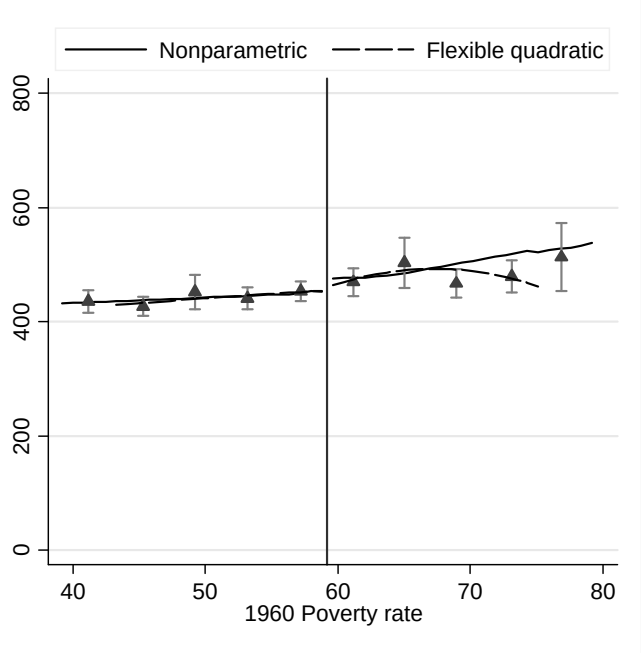


Figure 3: Estimated Discontinuity in Head Start Funding per 4 year old, 1968



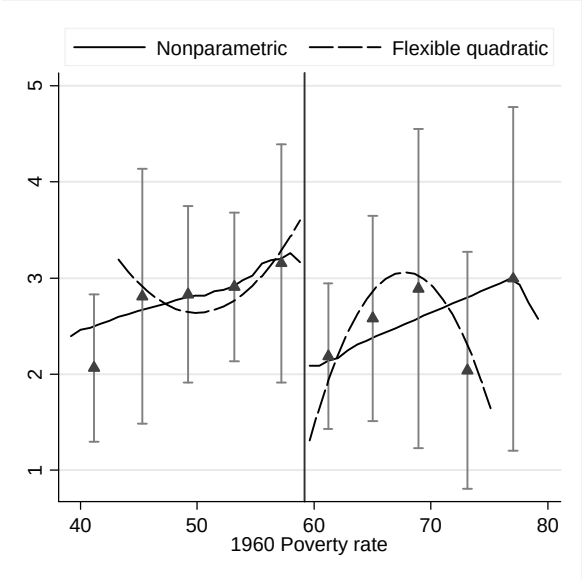
Source: Ludwig and Miller (2007)

Figure 4: Other Federal Per-Capita Social Spending, 1972 National Archives Data



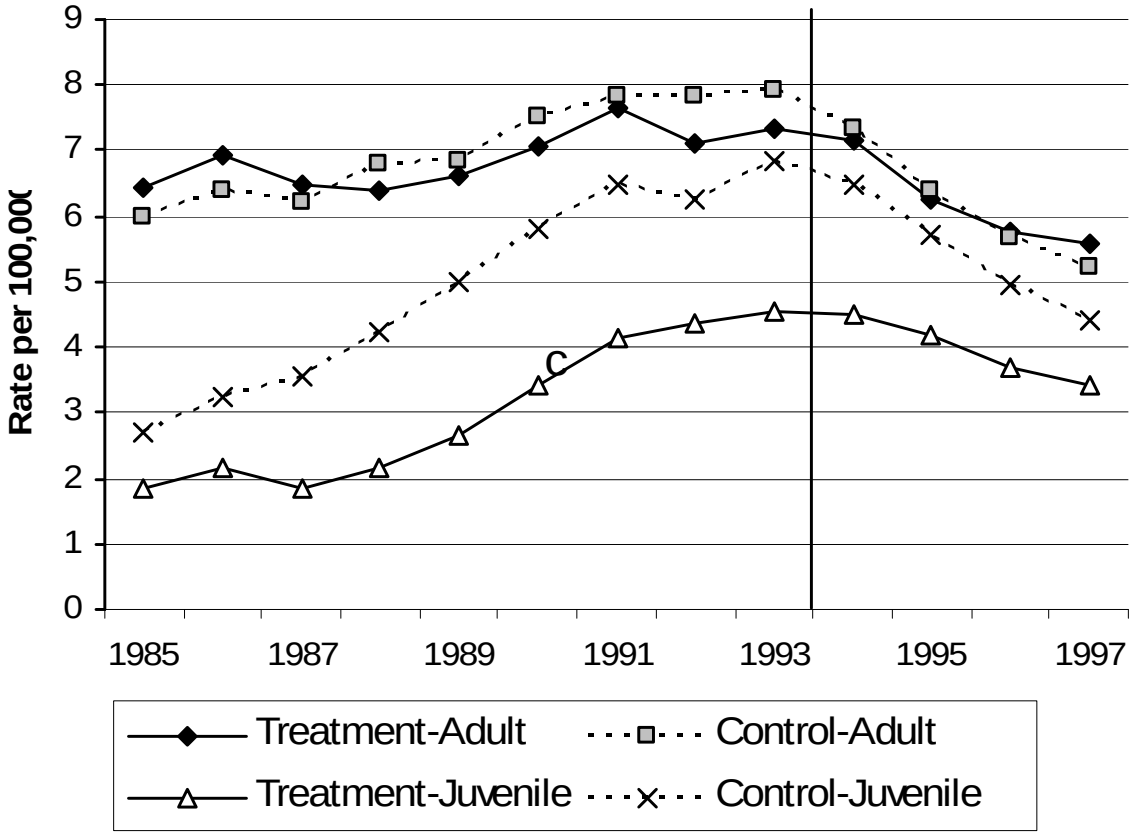
Source: Ludwig and Miller (2007)

Figure 5: Death rates to children 1973-83 from causes of death screened for by Head Start



Source: Ludwig and Miller (2007)

Figure 6. Gun homicide rates, Brady Treatment vs Control States, adults (over 21), and juveniles (under 21).



Source: Ludwig and Cook (2000).

Table 1:
Difference-in-difference analysis of effects of contaminated drinking water on cholera deaths, London

	1849	1852
Lambeth company	A	B
Southwark and Vauxhall company	D	D

Table 2: Social Costs of Crime from Ex Post and Ex Ante Perspectives (see text)

Source: Derived from Cohen (2005) and Miller, Cohen and Wiersema (1996), converted to 2009 Dollars.

	Productivity (ex ante)[11]	Medical care & ambulance (ex ante) [12]	Mental health care (ex ante)[13]	Police/fire services (ex ante)[14]	Social/victim services (ex ante)[15]	Property loss/damage (ex ante)[16]	Subtotal: tangible losses (ex ante)	Quality of Life (ex ante)[17]	Total Social costs of crime (ex ante method)	Willingness to Pay Estimate (ex post)
Fatal Crime/Homicide	$1,306,710	$23,734	$6,547	$1,791	$0	$14,286	$1,354,904	$2,671,162	$4,028,339	$10,641,866
Rape and sexual assault	$3,001	$682	$3,001	$50	$37	$136	$6,956	$111,029	$118,668	$260,013
Aggravated Assault	$4,228	$2,005	$132	$115	$63	$53	$6,547	$26,325	$32,736	$76,797
Simple Assault	$95	$0	$89	$94	$12	$20	$273	$2,319	$2,728	
Robbery or attempt	$1,296	$505	$90	$177	$34	$1,023	$3,137	$7,775	$10,912	$254,527
Drunk Driving	$3,819	$1,910	$112	$55		$2,182	$8,184	$16,232	$24,552	
Arson	$2,387	$1,500	$25	$1,364		$21,142	$26,598	$24,552	$51,150	
Larceny or attempt	$11	$0	$8	$109	$1	$368	$505	$0	$505	
Burglary or attempt	$16	$0	$7	$177	$7	$1,323	$1,500	$409	$1,910	$27,427
Motor vehicle theft or attempt	$61	$0	$7	$191	$0	$4,501	$4,774	$409	$5,047	

[11] Includes estimates of wages, fringe benefits, housework, and school days lost by victims and their families, productivity lost by co-workers and supervisors recruiting and training replacements for disabled workers, and insurance claims, processing costs, and legal expenses incurred in recovering productivity losses from drunk drivers and insurers. .

[12] Includes estimates of payments for hospital and physician care, emergency medical transport, rehabilitation, prescriptions, allied health services, medical devices, coroner costs, premature funeral expenses, and related insurance claims processing costs. Also included in this category are victim legal expenses incurred in recovering medical costs from drunk drivers and their insurers.

[13] Includes payments for services to crime victims by psychiatrists, psychologists, social workers, and pastoral counselors, and associated insurance claims processing costs.

[14] Includes initial police response and follow-up investigation, as well as fire service costs related to arson and drunk driving crashes. The cost of other aspects of the criminal justice system is not included.

[15] Estimated costs of activities of Victim Services Agencies and Child Protective Services agencies, as well as foster care for maltreated children removed from their homes, special education for maltreated children, and services aimed at reintegrating families with maltreatment problems.

[16] Includes the value of property damaged and of property taken and not recovered, plus insurance claims administration costs that arise in compensating victims' property losses.

[17] The nonfatal quality of life estimates come from regression analysis of jury verdicts. For nonfatal injuries, the standard error range around the log-linear regression estimates is +39 percent for assault and +29 percent for rape. Translating back to linear estimates of the quality of life yields extremely large confidence intervals.

Table 3: Cost Savings per Additional High School Graduate

	White Male	Black Male	White Female	Black Female	Average
Annual Earnings difference between High School Graduate and Drop-out	$11,100	$8,300	$8,700	$4,200	$8,075
Present Value of Lifetime Income and other Tax Benefits, per expected graduate	$202,700	$157,600	$109,100	$94,300	$139,100
Present Value of Lifetime Public Health Savings	$27,900	$52,100	$39,600	$62,700	$40,500
Present Value of lifetime cost-savings from reduced criminal activity	$30,200	$55,500	$8,300	$8,600	$26,600
Present Value of Welfare cost-saving per expected high school graduate	$1,200	$3,300	$5,000	$9,000	$3,000

Notes:
Source: Levin et al. (2007)

Discount rate used is 3.5%.

Tax benefits included are sales and property taxes, calculated by adding 5% to total income tax payments.

Public Health Savings calculated based on receiving Medicare and Medicaid payments

Criminal activity costs include estimates only for the costs of murder, rape/sexual assault, violent crime, property crime, and drugs offenses. Costs include specifically Bureau of Justice Statistics data and survey information which calculate the public cost per crime and per arrest for each of these five crime types. Each crime imposes costs in terms of policing, government programs to combat crime, and state-funded victim costs, trials, sentencing and incarceration.

Welfare Cost Savings includes estimates for reduced reliance on TANF, housing assistance and food stamps

Table 4: Average Annual and Present Value of Lifetime Earnings by Race and Education Level

Source: Average Annual Earnings from Levin et al. (2007), PV of lifetime earnings from U.S. Census Bureau Current Population Survey March 1998, 1999, 2000.

	Average Annual Earnings				PV of lifetime Earnings (millions)			
	High School drop out	High School Graduate	Some College	BA or higher	High School drop out	High School Graduate	Some College	BA or higher
White Male	$22,800	$33,900	$40,300	$79,100				
Black Male	$13,500	$21,800	$29,600	$53,800				
White Female	$7,800	$16,500	$20,400	$35,600				
Black Female	$10,000	$14,200	$19,500	$40,600				
White Average	$15,300	$25,200	$30,350	$57,350	$1.1	$1.3	$1.6	$2.2
Black Average	$11,750	$18,000	$24,550	$47,200	$.8	$1.0	$1.2	$1.7