

Return on Investment: Evidence-Based Options to Improve Statewide Outcomes

Technical Appendix Methods and User-Manual

This technical appendix describes the sources, assumptions, and computational methods used in the Washington State Institute for Public Policy's benefit-cost model.

**Stephanie Lee
Steve Aos
Elizabeth Drake
Annie Pennucci
Marna Miller
Laurie Anderson
Mason Burley**

April 9, 2012

Washington State Institute for Public Policy
110 Fifth Avenue Southeast, Suite 214 • PO Box 40999 • Olympia, WA 98504-0999
(360) 586-2677 • FAX (360) 586-2793 • www.wsipp.wa.gov

The authors would like to thank the considerable efforts of a number of people:

Roxanne Lieb, WSIPP

Corey Nunlist, WSIPP

Janie Maki, WSIPP

Laura Harmon, WSIPP

Tali Klima

Jim Mayfield

Margaret Kuklinski, Social Development Research Group

Gretchen Bjornstad, The Social Research Unit

Triin Edovald, The Social Research Unit

The Washington Legislature created the Washington State Institute for Public Policy in 1983. The Institute is governed by a Board of Directors that represents the legislature, governor, and public universities. The Board guides the development of all Institute activities. The mission of the Institute is to assist policymakers, particularly those in the legislature, in making informed judgments about important, long-term issues facing Washington State.

The Legislature authorized the Institute to receive outside funding for this project; the MacArthur Foundation supported 80 percent of the work and the Legislature funded the other 20 percent.

Washington State Institute for Public Policy

110 Fifth Avenue Southeast, Suite 214 • PO Box 40999 • Olympia, WA 98504-0999
(360) 586-2677 • FAX (360) 586-2793 • www.wsipp.wa.gov

Document No. 11-07-1201B

Contents

| | |
|---|-----------|
| Appendix A: Overview of the Benefit-Cost Approach and Model | 5 |
| A1. Overview of the Model | 6 |
| A2. General Approach and Characteristics of the Institute's Benefit-Cost Modeling Process | 6 |
| Appendix B: Methods Used to Estimate Effect Sizes and Standard Errors | 9 |
| B1. Meta-Analytic Procedures to Compute Effect Sizes and Standard Errors | 9 |
| B2. Procedures for Calculating Effect Sizes | 11 |
| B3. Institute Adjustments to Effect Sizes for Methodological Quality, Outcome Measure Relevance, Researcher Involvement, and Laboratory or Unusual Settings | 14 |
| B3.1 Methodological Quality | 15 |
| B3.2 Adjusting Effect Sizes for Research Involvement in the Program's Design and Implementation and for Laboratory or Unusual Settings..... | 16 |
| B3.3 Adjusting Effect Sizes for Evaluations With Weak Outcome Measures..... | 16 |
| B3.4 Values of Adjustment Factors..... | 17 |
| Appendix C: Procedures to Compute "Monetizable" Outcome Units | 18 |
| C1. Input Screen for Program Effect Size Parameters | 18 |
| C2. Unit Changes from Direct Effect Sizes | 20 |
| C3. Unit Changes From Linked Effect Sizes | 20 |
| C4. Monetizable Units for Benefit-Cost Calculation | 20 |
| Appendix D: Methods Used to Estimate Monetary Benefits | 21 |
| D1. Valuation of Outcomes That Affect Labor Market Earnings | 21 |
| D1.1 Earnings Data and Related Parameters..... | 21 |
| D2. Valuation of Outcomes That Affect Crime | 26 |
| D2.1 Per-Unit Crime Costs..... | 26 |
| D2.2 Criminological Information for Different Populations | 48 |
| D2.3 Estimates of Victimitizations Per Conviction | 52 |
| D2.4 Procedures to Estimate Criminal Justice System and Victimitization Events | 53 |
| D3. Valuation of Child Abuse and Neglect Outcomes..... | 58 |
| D3.1 Input Screens for CAN Parameters..... | 59 |
| D4. Valuation of Outcomes That Affect Alcohol and Illicit Drug Disorders, and Regular Tobacco Use..... | 66 |
| D4.1 Input Screens for ATOD Parameters | 67 |
| D4.2 ATOD Epidemiological Parameters: Current Prevalence for Prevention and Intervention Programs | 71 |
| D4.3 ATOD Attributable Deaths | 74 |
| D4.4 Linkages: ATOD and Other Outcomes..... | 76 |
| D4.5 Human Capital Outcomes Affecting Labor Market Earnings via ATOD-Caused Morbidity and Mortality | 76 |
| D4.6 Medical Costs, Treatment Costs, and Other Costs From ATOD | 77 |
| D4.7 Age of Initiation of ATOD..... | 79 |
| D5. Valuation of Teen Birth Outcomes..... | 80 |
| D6. Valuation of Public Assistance Outcomes | 80 |
| D7. Model Inputs for K-12 Education Outcomes..... | 82 |
| D7.1 Input Screens for Education Parameters | 82 |
| D7.2 Valuation of Earnings From High School Graduation | 84 |
| D7.3 Valuation of Earnings From Increases in K-12 Standardized Student Test Scores..... | 85 |
| D7.4 Valuation of Earnings from Increases in the Number of Years of Education Achieved .. | 86 |

| | | |
|---|---|------------|
| D7.5 | Valuation of Changes in the Use of K–12 Special Education and Grade Retention..... | 86 |
| D7.6 | Discount Factors for Decaying Test Score Effect Sizes to Age 17 | 87 |
| D8. | Valuation of Mental Health Outcomes | 88 |
| D8.1. | Input Screens for Mental Health Parameters. | 88 |
| D8.2. | Mental Health Epidemiological Parameters | 94 |
| D8.3. | Linkages: Mental Health to Other Outcomes | 95 |
| D8.4. | Human Capital Outcomes Affecting Labor Market Earnings via Mental Health Morbidity and Mortality | 95 |
| D8.5 | Medical Costs | 97 |
| D9. | Health Care Parameters | 98 |
| D10. | Other Parameters | 100 |
| D10.1 | Base Year for Monetary Denomination | 100 |
| D10.2 | Discount Rates | 101 |
| D10.3 | Demographic Information | 102 |
| D10.4 | Valuation of Reductions in Mortality Risk: Value of a Statistical Life | 103 |
| D10.5 | Deadweight Cost of Taxation | 106 |
| D10.6 | Inflation/Price Indexes | 107 |
| D10.7 | Household Production | 108 |
| D10.8 | Tax Rates | 110 |
| D10.9 | Capital Costs | 110 |
| Appendix E: Meta Analyses of Linked Outcomes | | 111 |
| E1. | Input Screen for Linked Outcome Effect Sizes | 111 |
| E2. | Institute Adjustments to Effect Sizes for Methodological Quality, Generalizability of the Sample, and Relevance of the Independent and Dependent Variables | 112 |
| E2.1 | Methodological Quality | 112 |
| E2.2 | Generalizability of the Sample..... | 112 |
| E2.3 | Relevance of the Independent and Dependent Variables..... | 112 |
| Appendix F: Methods to Access Risk and Uncertainty | | 126 |
| F1. | Key Inputs Varied in the Monte Carlo Simulation Analysis | 126 |
| F2. | Computational Procedures to Carry Out the Simulation | 127 |
| Appendix G: The WSIPP Sentencing and Programming Portfolio Tool..... | | 129 |
| G1. | Estimating the Crime Effects of Sentencing-Related Policies (Step 1 of 5)..... | 131 |
| G2. | Estimating the Fiscal Effect of Sentencing-Related Policy Changes (Step 2 of 5) | 138 |
| G3. | Estimating the Effect of a Portfolio of Programs That Reduce Crime (Step 3 of 5) | 140 |
| G4. | Combinations of Policies (Step 4 of 5) | 142 |
| G5. | Risk Analysis (Step 5 of 5)..... | 143 |

Appendix A: Overview of the Benefit-Cost Approach and Model

This technical appendix describes the latest version of the Washington State Institute for Public Policy (Institute) benefit-cost model. The Institute built its first model in 1997 to estimate the economic value of programs that reduce crime. Later, as the Institute received additional and varied assignments from the Washington legislature, the benefit-cost model was revised and expanded to cover additional public policy outcomes. The model described here reflects our current approach to computing benefits and costs for a wide array of outcomes and contains several enhancements over earlier versions. Our ongoing goal is to provide Washington policy makers with better “bottom-line” estimates each successive legislative session.

The 2009 Washington State Legislature directed to update and extend its review of the benefits and costs of prevention and intervention programs.¹ The Legislature directed the Institute to “calculate the return on investment to taxpayers from evidence-based prevention and intervention programs and policies.” Specifically, the Legislature asked the Institute to identify public policies that have been shown to improve these broad outcomes of public interest:

- Crime,
- K–12 education,
- Child maltreatment,
- Substance abuse,
- Mental health,
- Public health,
- Public assistance,
- Employment, and
- Housing.

A principal objective of the Institute’s model is to produce a “What Works?” list of public policy options available to the Washington State legislature—and to rank the list by estimates of return on investment. The ranked list can then help policy makers choose a portfolio of public policies that are evidence based and that have a high likelihood of producing more benefits than costs. For example, if the public policy objective is to reduce crime, then a portfolio of evidence-based policies can be selected from the list—from prevention policies, juvenile justice policies, and adult corrections policies—that together can improve the chance that crime is reduced and taxpayer money is used efficiently.

There are three basic steps to the analysis.

1. **What Works?** First, we conduct a systematic review of the research literature to identify policies and programs that have demonstrated an ability to improve the outcomes. In Appendices B and C, we describe the methods we use to screen and code research studies, the meta-analytic approach we use to estimate the effectiveness of policy options to achieve outcomes, and the procedures we use to compute monetizable units of change. The objective of the first step is to draw statistical conclusions about what works—and what does not—to achieve improvements in the outcomes, along with an estimate of the statistical error involved.
2. **What Makes Economic Sense?** The second basic step involves applying economic calculations to put a monetary value on the improved outcomes (from the first step). Once monetized, the estimated benefits are then compared to the costs of programs to arrive at a set of economic bottom lines for the investments. Appendix D describes the processes we use to monetize the outcomes.
3. **How Risky are the Estimates?** Part of the process of estimating a return on investment involves assessing the riskiness of the estimates. Any rigorous modeling process, such as the one described here, involves many individual estimates and assumptions. Almost every step involves at least some level of uncertainty. Appendix F describes the “Monte Carlo” approach we use to model this uncertainty. The objective of the risk analysis is to assess the odds that an individual return on investment estimate may offer the legislature the wrong advice. For example, if we conclude that, on average, an investment in program XYZ has a ratio of three dollars of benefits for each dollar of cost, what are the odds, given the uncertainty in this estimate, that the program will not even generate one dollar of benefits for each dollar of cost?

Thus, our analytical goal for each evidence-based investment option we analyze is to deliver to the legislature two benefit-cost bottom-line measures: an expected return on investment and, given the uncertainty, the odds that the investment will at least break even.

¹ Laws of 2009, Ch. 564, § 610 (4).

A1. Overview of the Model

The Institute's benefit-cost model is an integrated set of estimates and computational routines designed to produce four related benefit-cost summary statistics: net present value, benefit-to-cost ratio, internal rate of return on investment, and measure of risk associated with these bottom-line estimates. In simplest form, the model implements a standard economic calculation of the expected worth of an investment by computing the net present value of a stream of estimated benefits and costs that occur over time, as described with equation (A1).

$$(A1) \quad NPV_{tage} = \sum_{y=tage}^N \frac{Q_y \times P_y - C_y}{(1 + Dis)^y}$$

In this basic model, the net present value, NPV , of a program is the quantity of the outcomes achieved by the program or policy, Q , in year y , times the price per unit of the outcome, P , in year y , minus the cost of producing the outcome, C , in year y . The lifecycle of each of these values is measured from the average age of the person who is treated, $tage$, and runs over the number of years into the future over which they are evaluated, N . The future values are expressed in present value terms after applying a discount rate, Dis .

The first term in the numerator of equation (A1), Q_y , is the estimated number of outcome "units" in year y produced by the program or policy. The procedures used to develop estimates of Q_y are described in Appendices B and C. In Appendix D we describe the various methods we use to estimate the price term, P_y , in equation (A1). In Appendix F, we describe the Monte Carlo simulation procedures we employ to estimate the risk and uncertainty in the single-point net present value estimates.

Rearranging terms in (A1), a benefit-to-cost ratio, B/C , can be computed with:

$$(A2) \quad \frac{B}{C} = \frac{\sum_{y=tage}^N \frac{Q_y \times P_y}{(1 + Dis)^y}}{\sum_{y=tage}^N \frac{C_y}{(1 + Dis)^y}}$$

Additionally, since the model keeps track of the estimated annual cash flows of benefits and costs of a program, an internal rate of return on investment can be computed. The internal rate of return is the discount rate, in equation (A1), that results in a zero net present value. In computations, the internal rate of return is calculated using Microsoft Excel's[®] IRR function. For some cash flow series, internal rates of return cannot be calculated.

A2. General Approach and Characteristics of the Institute's Benefit-Cost Modeling Process

There are several features that are central to the Institute's benefit-cost modeling approach.

Internal Consistency of Estimates. Because the Institute's model is used to evaluate the benefits and costs of a wide range of public policies that affect many different outcomes, a key modeling goal is internal consistency. Any complex investment analysis, whether geared toward private sector or public sector investments, involves many estimates and uncertainties. Across all the outcomes and programs we consider, we attempt to be as internally consistent as possible. That is, within each topic area, our bottom-line estimates are developed so that a net present value for one program can be compared directly to that of another program. This is in contrast to the way most benefit-cost analyses are done, where one study conducts an economic analysis for one program and then another study performs a different benefit-cost analysis for another program—the result can often lead to apples and oranges comparisons. By adopting one modeling approach to assess all decisions, on the other hand, the consistency of results is enhanced, thereby enabling apples-to-apples benefit-to-cost comparisons.

Meta analytic strategy. The first step in our benefit-cost modeling strategy produces estimates of policies and programs that have been shown to improve particular outcomes. We carefully analyze all high-quality studies from the United States and elsewhere to identify well-researched interventions that have achieved outcomes (as well as those that have not). We look for research studies with strong, credible evaluation designs, and we ignore studies with weak research methods. Our empirical approach follows a meta-analytic framework to assess systematically all relevant evaluations we can locate on a given topic. We focus the topics on those policies or programs that are the subject of budget or policy decisions facing the Washington legislature. By including all of the studies in a meta-analysis, we are, in effect, making an average statement about the effectiveness of all relevant studies on a particular topic. For example, in deciding whether the program "Functional Family Therapy" works to reduce crime, we do not rely on just one evaluation of the program. Rather, we compute a meta-analytic average effect from all eight credible studies we find of Functional Family Therapy.

Long-Run Benefits and Costs. We include estimates of the long-term benefits and costs of programs and policies. In most cases, this involves Institute projections well into the future. Projections are necessary, because most of the evidence about programs comes from evaluations with relatively short follow-up periods. It is rare to find longitudinal program evaluations. This problem, of course, is not unique to public programs. Most private investment decisions are based on past performance, and future results are projected by entrepreneurs or investment advisors based on certain assumptions. We adopt that private-sector investment approach in this model. We forecast, using a consistent and empirically based framework, the long-term benefits and costs of programs and policies. We then assess the riskiness of the projections.

Sums of Benefits. Many evaluations of programs and policies measure multiple outcomes. In most cases, we sum the per-participant benefits across multiple outcomes to draw a comprehensive conclusion about the total benefits to society. However, some categories of benefits draw from the same ultimate monetary source. For example, high school graduation and standardized test scores are two outcomes that may both be measured by a program evaluation. We have methods to monetize both outcomes, but improvements in both lead to increased earnings in the labor market. To avoid “double-counting,” rather than summing the benefits from both high school graduation and standardized test scores, we compute a weighted average of the two, based on the total effective sample sizes in the research studies that measured each outcome.

Risk. Any tabulation of benefits and costs necessarily involves uncertainty and some degree of speculation about future performance. This is expected in any investment analysis. Therefore, it is important to understand how conclusions might change when assumptions are altered. To assess risk, we perform a “Monte Carlo simulation” technique in which we vary the key factors in our calculations. The purpose of the risk analysis is to determine the odds that a particular approach will at least break-even. We are interested in the expected rate of return on investment for any program, but we also want to calculate the odds that a particular program will not break even. This type of risk and uncertainty analysis is used by many businesses in investment decision making; we employ the same tools to test the riskiness of the public sector options considered in this report

Three Perspectives on Benefits and Costs. We present these monetary estimates from three distinct perspectives: the benefits that accrue solely to program participants, those received by taxpayers, and any other measurable (non-participant and non-taxpayer) monetary benefits.

The sum of these three perspectives provides a “total Washington” view on whether a program produces benefits that exceed costs. Restricting the focus solely to the taxpayer perspective can also be useful for certain fiscal analysis and state budget preparation.

For example, we estimate the long-term labor market benefits that accrue directly to the participants in a successful early childhood education program. As we show in this analysis, there is also evidence that a successful early childhood education program produces lower long-term crime rates and, thus, generates benefits to non-participants by lowering the amount of money that taxpayers have to spend on the criminal justice system. Lower crime also reduces the amount of costs that crime victims would otherwise have to bear. Thus, we provide estimates for each of the three perspectives: program participants, non-participants as taxpayers, and non-participants in other non-taxpayer roles.

The Model’s Expandability. The state of research-based knowledge is continually expanding. More is known today than ten years ago on the relative effectiveness of prevention and intervention programs, and more will be known in the future. We built this benefit-cost model so that it can be expanded to incorporate this evolving state of research-based evidence. Similar to an investment analyst’s model used to update quarterly earnings-per-share estimates of private investments, this model is designed to be updated regularly as new and better information becomes available. This design feature allows increasingly refined estimates of the economic bottom lines for prevention and early intervention programs, and will supply government decision makers with the latest information on how taxpayers can get better returns on their dollars.

New Features in the Current Version. As noted, we have been developing the general benefit-cost model for over a decade. In the current update, we have added a number of new features. Some of the major ones include:

- **Software user-interface.** As shown in this appendix, we have developed software to enable users, including the Institute, to enter and store information and run programs in a more user-friendly manner. All of the input assumptions and estimates are included on user-input forms. This appendix describes the current values for each input.
- **Monte Carlo simulation.** The current version of the model implements a consistent risk analysis for key inputs, see Appendix F.

- **Deadweight cost of taxation.** Following the work of Heckman et al.², we have implemented a deadweight cost of taxation for any taxpayer cost or benefit estimated with the model.
- **Value of reduction in mortality: Value of a Statistical Life (VSL).** For any outcomes that affect mortality, we now include user-inputs for the value of a statistical life. We continue to model mortality-related lost earnings and household production separately, and adjust the VSL estimate to be a “net-VSL” estimate (after deducting the individually measured earnings and household production estimates).
- **Children’s mental health.** The model now explicitly measures the value of changes in certain child internalizing and externalizing mental health outcomes.
- **Medical expenditures.** This version of the model uses more direct modeling of health care costs.

² Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1-2), 114-128.

Appendix B: Methods Used to Estimate Effect Sizes and Standard Errors

As outlined in Appendix A, the Institute model is an integrated set of estimates and computational routines designed to produce internally consistent benefit-to-cost estimates for a variety of public policies and programs. The model implements a standard economic calculation of the expected worth of an investment by computing the net present value of a stream of estimated benefits and costs that occur over time, as described with equation (B).

$$(B) \quad NPV_{tage} = \sum_{y=tage}^N \frac{Q_y \times P_y - C_y}{(1 + Dis)^y}$$

In this basic model, the net present value, NPV , of a program is the quantity of the outcomes produced by the program or policy, Q , in year y , times the price per unit of the outcome, P , in year y , minus the cost of producing the outcome, C , in year y . The lifecycle of each of these values is measured from the average age of the person treated, $tage$ and runs over the number of years into the future over which they are evaluated, N . The future values are expressed in present value terms after applying a discount rate, Dis .

The first term in the numerator of equation (B), Q_y , is the estimated number of outcome “units” in year y produced by the program or policy. The procedures used to develop estimates of Q_y are described in Appendices B and C. In Appendix D we describe the various methods we use to estimate the price term, P_y , in equation (B).

This appendix describes the process we use to estimate effect sizes—the central element of Q_y , in equation (B).

B1. Meta-Analytic Procedures to Compute Effect Sizes and Standard Errors

To estimate the effects of programs and policies on outcomes, we employ statistical procedures researchers have been developing to facilitate systematic reviews of evaluation evidence. This set of procedures is called “meta-analysis” and we employ that methodology in this study.³

Study Selection and Coding Criteria

A meta-analysis is only as good as the selection and coding criteria used to conduct the study.⁴ Following are the key choices we made and implemented.

Study Selection. We used four primary means to locate studies for meta-analysis of programs: (1) we consulted the bibliographies of systematic and narrative reviews of the research literature in the various topic areas; (2) we examined the citations in the individual studies themselves; (3) we conducted independent literature searches of research databases using search engines such as Google, Proquest, Ebsco, ERIC, PubMed, and SAGE; and (4) we contacted authors of primary research to learn about ongoing or unpublished evaluation work. As we will describe, the most important criteria for inclusion in our study was that an evaluation have a control or comparison group. Therefore, after first identifying all possible studies via these search methods, we attempted to determine whether the study was an outcome evaluation that had a comparison group. We also determined if each study used outcome measures that were standardized or well-validated. If a study met these criteria, we then secured a paper copy of the study for our review.

Peer-Reviewed and Other Studies. We examined all evaluation studies we could locate with these search procedures. Many of these studies were published in peer-reviewed academic journals while many others were from reports obtained from the agencies themselves. It is important to include non-peer reviewed studies, because it has been suggested that peer-reviewed publications may be biased to show positive program effects. Therefore, our meta-analysis includes all available studies that meet our other criteria, regardless of published source.

Control and Comparison Group Studies. Our analysis only includes studies that had a control or comparison group. That is, we did not include studies with a single-group, pre-post research design. This choice was made because it is only through rigorous comparison group studies that causal relationships can be reliably estimated.

³ In general, we follow the meta-analytic methods described in: Lipsey, M. W., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

⁴ All studies used in the meta-analysis are identified in the references in Technical Appendix 1. Many other studies were reviewed, but did not meet the criteria set for this analysis.

Exclusion of Studies of Program Completers Only. We did not include a study in our meta-analytic review if the treatment group was made up solely of program completers. We adopted this rule because there are too many significant unobserved self-selection factors that distinguish a program completer from a program dropout, and these unobserved factors are likely to significantly bias estimated treatment effects. Some studies of program completers, however, also contain information on program dropouts in addition to a comparison group. In these situations, we included the study if sufficient information was provided to allow us to reconstruct an intent-to-treat group that included both completers and non-completers, or if the demonstrated rate of program non-completion was very small. In these cases, the study still needed to meet the other inclusion requirements listed here.

Random Assignment and Quasi-Experiments. Random assignment studies were preferred for inclusion in our review, but we also included non-randomly assigned comparison groups. We only included quasi-experimental studies if sufficient information was provided to demonstrate comparability between the treatment and comparison groups on important pre-existing conditions such as age, gender, and pre-treatment characteristics such as test scores or level of functioning.

Enough Information to Calculate an Effect Size. Following the statistical procedures in Lipsey and Wilson,⁵ a study had to provide the necessary information to calculate an effect size. If the necessary information was not provided, and we were unable to obtain the necessary information directly from the study author(s), the study was not included in our review.

Mean-Difference Effect Sizes. For this study, we coded mean-difference effect sizes following the procedures in Lipsey and Wilson.⁶ For dichotomous measures, we used the D-cox transformation to approximate the mean difference effect size, as described in Sánchez-Meca, Marín-Martínez, and Chacón-Moscoso.⁷ We chose to use the mean-difference effect size rather than the odds ratio effect size because we frequently coded both dichotomous and continuous outcomes (odds ratio effect sizes could also have been used with appropriate transformations).

Multivariate Results Preferred. Some studies presented two types of analyses: raw outcomes that were not adjusted for covariates such as age, gender, or pre-intervention characteristics; and those that had been adjusted with multivariate statistical methods. In these situations, we coded the multivariate outcomes and used test statistics from the regression to calculate an effect size.

Outcome Measures of Interest. Our primary outcomes of interest include standardized, validated measurements. A list of the outcomes coded in the program areas are listed in Technical Appendix 1. Where possible, our model estimates monetary values for the outcomes. At this time, however, we are not able to monetize all of these outcomes.

Averaging Effect Sizes for Similar Outcomes. Some studies reported similar outcomes: e.g., arrest and incarceration, or reading test scores from different standardized assessments. In such cases, we meta-analyzed the similar measures and used the combined effect size in the meta-analysis for that program. As a result, each study sample coded in this analysis is associated with a single effect size for a given outcome.

Dichotomous Measures Preferred Over Continuous Measures. Some studies included two types of measures for the same outcome: a dichotomous (yes/no) outcome and a continuous (mean number) measure. In these situations, we coded an effect size for the dichotomous measure. Our rationale for this choice is that in small or relatively small sample of studies, continuous measures of treatment outcomes can be unduly influenced by a small number of outliers, while dichotomous measures can avoid this problem. Of course, if a study only presented a continuous measure, we coded the continuous measure.

Longest Follow-Up Periods. When a study presented outcomes with varying follow-up periods, we generally coded the effect size for the longest follow-up period. The longest follow-up period allows us to gain the most insight into the long-run benefits and costs of various treatments. Occasionally, we did not use the longest follow-up period if it was clear that a longer reported follow-up period adversely affected the attrition rate of the treatment and comparison group samples.

If outcomes for study samples are measured at multiple points in time, and if a sufficient number of studies contained multiple, similar follow-up periods, we calculated effect sizes for an initial and longer term follow-up period. Using different points of time of measurement allows us to examine whether program effects “fade out” over time.

Some Special Coding Rules for Effect Sizes. Most studies in our review had sufficient information to code exact mean-difference effect sizes. Some studies, however, reported some, but not all the information required. We followed the following rules for these situations:

- **Two-tail p-values.** Some studies only reported p-values for significance testing of program outcomes. When we had to rely on these results, if the study reported a one-tail p-value, we converted it to a two-tail test.

⁵ Lipsey & Wilson, 2001.

⁶ Ibid.

⁷ Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8(4), 448-467.

- **Declaration of significance by category.** Some studies reported results of statistical significance tests in terms of categories of p-values. Examples include: $p \leq .01$, $p \leq .05$, or non-significant at the $p = .05$ level. We calculated effect sizes for these categories by using the highest p-value in the category. Thus, if a study reported significance at $p \leq .05$, we calculated the effect size at $p = .05$. This is the most conservative strategy. If the study simply stated a result was non-significant, we computed the effect size assuming a p-value of .50.

B2. Procedures for Calculating Effect Sizes

Effect sizes summarize the degree to which a program or policy affects an outcome. In experimental settings this involves comparing the outcomes of treated participants relative to untreated participants. There are several methods used by analysts to calculate effect sizes, as described in Lipsey and Wilson.⁸ The most common effect size statistic is the standardized mean difference effect size, and that is the measure we employ in this analysis.

Continuously Measured Outcomes. The mean difference effect size was designed to accommodate continuous outcome data, such as student test scores, where the differences are in the means of the outcome.⁹ The standardized mean difference effect size is computed with:

$$(B1) ES = \frac{M_t - M_c}{\sqrt{\frac{(N_t - 1)SD_t^2 + (N_c - 1)SD_c^2}{N_t + N_c - 2}}}$$

In this formula, ES is the estimated effect size for a particular program; M_t is the mean value of an outcome for the treatment or experimental group; M_c is the mean value of an outcome for the control group; SD_t is the standard deviation of the treatment group; and SD_c is the standard deviation of the control group; N_t is the number of subjects in the treatment group; and N_c is the number of subjects in the control group.

The variance of the mean difference effect size statistic in (B1) is computed with:¹⁰

$$(B2) ESVar = \frac{N_t + N_c}{N_t N_c} + \frac{ES^2}{2(N_t + N_c)}$$

In some random assignment studies or studies where treatment and comparison groups are well-matched, authors provide only statistical results from a t-test. In those cases, we calculate the mean difference effect size using:¹¹

$$(B3) ES = t \sqrt{\frac{N_t + N_c}{N_t N_c}}$$

In many research studies, the numerator in (B1), $M_t - M_c$, is obtained from a coefficient in a regression equation, not from experimental studies of separate treatment and control groups. For such studies, the denominator in (B1) is the standard deviation for the entire sample. In these types of regression studies, unless information is presented that allows the number of subjects in the treatment condition to be separated from the total number in a regression analysis, the total N from the regression is used for the sum of N_t and N_c , and the product term $N_t N_c$ is set to equal $(N/2)^2$.

Dichotomously Measured Outcomes. Many studies record outcomes not as continuous measures such as test scores, but as dichotomies; for example, high school graduation. For these yes/no outcomes, Sanchez-Meca, et al.¹² have shown that the Cox transformation produces the most unbiased approximation of the standardized mean effect size. Therefore, to approximate the standardized mean difference effect size for continuously measured outcomes, we calculate the effect size for dichotomously measured outcomes with:

$$(B4) ES_{Cox} = \frac{\ln \left[\frac{P_t(1 - P_c)}{P_c(1 - P_t)} \right]}{1.65}$$

⁸ Lipsey & Wilson, 2001.

⁹ Ibid, Table B10, equation 1, p. 198.

¹⁰ Ibid, Table 3.2, p. 72.

¹¹ Ibid, Table B10, equation 2, p. 198

¹² Sanchez-Meca et al., 2003.

where \bar{O}_t is the percentage of the treatment group with the outcome and \bar{O}_c is the percentage of the comparison group with the outcome. The numerator, the logged odds ratio, is then divided by 1.65.

The ES_{Cox} has a variance of

$$(B5) \quad ESVar_{Cox} = .367 \left[\frac{1}{O_{1t}} + \frac{1}{O_{2t}} + \frac{1}{O_{1c}} + \frac{1}{O_{2c}} \right]$$

where O_{1t} , O_{2t} , O_{1c} , and O_{2c} are the number of successes (1) and failures (2) in the treatment, t, and control, c groups.

Occasionally when outcomes are dichotomous, authors reported the results of statistical analysis such as Chi-Square (X^2) statistics. In these cases, we first estimate the absolute value of $ES_{arcsine}$ per Lipsey and Wilson¹³, then based on analysis we conducted, we multiply the result by 1.35 to determine ES_{Cox} .

$$(B6) \quad |ES_{Cox}| = 1.35 * 2 \sqrt{\frac{X^2}{N_t + N_c - X^2}}$$

Similarly, we determined that in these cases, using (B2) to calculate the variance underestimates $ESVar_{Cox}$ and, hence over estimates the inverse variance weight. We conducted analysis which showed that $ESVar_{Cox}$ is linearly related to $ESVar$. Our analysis indicated that by multiplying $ESVar$ by 1.65 provides a very good approximation of $ESVar_{Cox}$.

Pre/Post Measures. Where authors report pre- and post-treatment measures without other statistical adjustments, first we calculate two between-groups effect sizes: (1) at pre-treatment and, (2) at post-treatment. Finally, we calculate the overall effect size by subtracting the post-treatment effect size from the pre-treatment effect size.

Adjusting Effect Sizes for Small Sample Sizes

Since some studies have very small sample sizes, we follow the recommendation of many meta-analysts and adjust for this. Small sample sizes have been shown to upwardly bias effect sizes, especially when samples are less than 20. Following Hedges,¹⁴ Lipsey and Wilson¹⁵ report the “Hedges correction factor,” which we use to adjust all mean-difference effect sizes, (where N is the total sample size of the combined treatment and comparison groups):

$$(B7) \quad ES'_m = \left[1 - \frac{3}{4N - 9} \right] * ES_m$$

Adjusting Effect Sizes and Variances for Multi-Level Data Structures. Most studies in the education field use data that are hierarchical in nature. That is, students are clustered in classrooms, classrooms are clustered within schools, schools are clustered within districts, and districts are clustered within states. Analyses that do not account for clustering will underestimate the variance in outcomes at the student level (the denominator in equation B1 and, thus, may over-estimate effect sizes. In studies that do not account for clustering, effect sizes and their variance require additional adjustments.¹⁶

There are two types of studies, each requiring a different set of adjustments.¹⁷

First, for student-level studies that ignore the variance due to clustering, we make adjustments to the mean effect size and its variance,

$$(B8) \quad ES_T = ES_m * \sqrt{1 - \frac{2(n-1)\rho}{N-2}}$$

¹³ Lipsey and Wilson, 2001, Table B10, equation 23, p. 200

¹⁴ Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107-128.

¹⁵ Lipsey & Wilson, 2001, equation 3.22, p. 49.

¹⁶ Studies that employ hierarchical linear modeling, or fixed effects with robust standard errors, or random effects models account for variance and need no further adjustment.

¹⁷ These formulas are taken from: Hedges, L. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, 32(4), 341-370.

$$(B9) V\{ES_T\} = \left(\frac{N_t + N_c}{N_t N_c} \right) [1 + (n - 1)\rho] + ES_T^2 \left(\frac{(N - 2)(1 - \rho)^2 + n(N - 2n)\rho^2 + 2(N - 2n)\rho(1 - \rho)}{2(N - 2)[(N - 2) - 2(n - 1)\rho]} \right)$$

where ρ is the intraclass correlation, the ratio of the variance between clusters to the total variance; N is the total number of individuals in the treatment group, N_t , and the comparison group, N_c ; and n is the average number of persons in a cluster, K .

In the educational field, clusters can be classes, schools, or districts. For this study, we used 2006 Washington Assessment of Student Learning (WASL) data to calculate values of ρ for the school-level ($\rho = 0.114$) and the district level ($\rho = 0.052$). Class-level data are not available for the WASL, so we use a value of $\rho = 0.200$ for class-level studies.

Second, for studies that report means and standard deviations at a cluster level, we make adjustments to the mean effect size and its variance:

$$(B10) ES_T = ES_m * \sqrt{\frac{1 + (n - 1)\rho}{n\rho}} * \sqrt{\rho}$$

$$(B11) v\{ES_T\} = \left\{ \left(\frac{N_t - N_c}{N_t N_c} \right) * \left(\frac{1 + (n - 1)\rho}{n\rho} \right) + \frac{[1 + (n - 1)\rho]^2 * ES_T^2}{2n\rho(K - 2)} \right\} * \rho$$

We did not adjust effect sizes in studies reporting dichotomous outcomes. This is because the Cox transformation assumes the entire normal distribution at the student level.¹⁸ However, when outcomes are dichotomous, or an effect size is calculated from studies where authors control for clustering with robust standard errors or hierarchical linear modeling, we use the “design effect” to calculate the “effective sample size”.¹⁹ The design effect is given by:

$$(B12) D = 1 + (n - 1)\rho$$

And the effective sample size is the actual sample size divided by the design effect. For example the effective sample size for the treatment group is:

$$(B13) N_{t(eff)} = \frac{N_t}{D}$$

Computing Weighted Average Effect Sizes, Confidence Intervals, and Homogeneity Tests. Once effect sizes are calculated for each program effect, and any necessary adjustments for clustering are made, the individual measures are summed to produce a weighted average effect size for a program area. We calculate the inverse variance weight for each program effect and these weights are used to compute the average. These calculations involve three steps. First, the standard error, SE_T of each mean effect size is computed with:²⁰

$$(B14) SE_T = \sqrt{\frac{N_t + N_c}{N_t N_c} + \frac{ES^2}{2(N_t + N_c)}}$$

Next, the inverse variance weight w is computed for each mean effect size with:²¹

$$(B15) w = \frac{1}{SE_T^2}$$

The weighted mean effect size for a group with i studies is computed with:²²

¹⁸ Mark Lipsey (personal communication, November 11, 2007).

¹⁹ Formulas for design effect and effective sample size were obtained from the Cochrane Reviewers Handbook, section 16.3.4,

Approximate analyses of cluster-randomized trials for a meta-analysis: effective sample sizes. <http://www.cochrane-handbook.org/>

²⁰ Lipsey & Wilson, 2001, equation 3.23, p. 49.

²¹ Ibid., equation 3.24, p. 49.

²² Ibid., p. 114

$$(B16) \overline{ES} = \frac{\sum(w_i ES_{T_i})}{\sum w_i}$$

Confidence intervals around this mean are then computed by first calculating the standard error of the mean with:²³

$$(B17) SE_{\overline{ES}} = \sqrt{\frac{1}{\sum w_i}}$$

Next, the lower, ES_L , and upper limits, ES_U , of the confidence interval are computed with:²⁴

$$(B18) \overline{ES}_L = \overline{ES} - z_{(1-\alpha)} (SE_{\overline{ES}})$$

$$(B19) \overline{ES}_U = \overline{ES} + z_{(1-\alpha)} (SE_{\overline{ES}})$$

In equations (B18) and (B19), $z_{(1-\alpha)}$ is the critical value for the z -distribution (1.96 for $\alpha = .05$).

The test for homogeneity, which provides a measure of the dispersion of the effect sizes around their mean, is given by:²⁵

$$(B20) Q_i = \left(\sum w_i ES_i^2 \right) - \frac{\left(\sum w_i ES_i \right)^2}{\sum w_i}$$

The Q-test is distributed as a chi-square with $k-1$ degrees of freedom (where k is the number of effect sizes).

Computing Random Effects Weighted Average Effect Sizes and Confidence Intervals. Next, a random effects model is used to calculate the weighted average effect size. Random effects models allow us to account for between-study variance in addition to within-study variance.²⁶

This is accomplished by first calculating the random effects variance component, v ²⁷

$$(B21) v = \frac{Q_i - (k - 1)}{\sum w_i - (\sum w_i^2 / \sum w_i)} \quad \text{Type equation here.}$$

where w_i is the square of the weight of ES_i (B15).

This random variance factor is then added to the variance of each effect size and finally all inverse variance weights are recomputed, as are the other meta-analytic test statistics. If the value of Q is less than the degrees of freedom ($k-1$), there is no excess variation between studies and the initial variance estimate is used.

B3. Institute Adjustments to Effect Sizes for Methodological Quality, Outcome Measure Relevance, Researcher Involvement, and Laboratory or Unusual Settings

In Technical Appendix 1, we show the results of our meta-analyses calculated with the standard meta-analytic formulas described in Appendix B2, above. In the last column of the exhibit, however, we list the “Adjusted Effect Size” that we actually use in our analysis. These adjusted effect sizes, which are derived from the unadjusted results, may be smaller, larger or equal to the unadjusted effect sizes we report in the same exhibit.

In this section, we describe our rationale for making these adjustments. In particular, we make four types of adjustments that are necessary to better estimate the results that we are more likely to achieve in real-world settings. We make

²³ Ibid.

²⁴ Ibid.

²⁵ Ibid., p. 116

²⁶ Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, 1(2), 97-111.

²⁷ Ibid., p. 134

adjustments for: (a) the methodological quality of each study we include in the meta-analyses; (b) the relevance or quality of the outcome measure that individual studies used; (c) the degree to which the researcher(s) who conducted a study were invested in the program's design; and (d) laboratory or other unusual, non-"real world" settings.

B3.1 Methodological Quality

Not all research is of equal quality, and this greatly influences the confidence that can be placed in the results of a study. Some studies are well-designed and implemented, and the results can be viewed as accurate representations of whether the program itself worked. Other studies are not designed as well, and less confidence can be placed in any reported differences. In particular, studies of inferior research design cannot completely control for sample selection bias or other unobserved threats to the validity of reported research results. This does not mean that results from these studies are of no value, but it does mean that less confidence can be placed in any cause-and-effect conclusions drawn from the results.

To account for the differences in the quality of research designs, we use a 6-point scale (with values ranging from zero to five) as a way to adjust the reported results. On this scale, a rating of "5" reflects an evaluation in which the most confidence can be placed: a well-implemented random assignment study. Generally, as the evaluation ranking gets lower, less confidence can be placed in any reported differences (or lack of differences) between the program and comparison or control groups.²⁸ A rating of "0" reflects an evaluation that does not have a comparison group or has a comparison group that is not equivalent to the treatment group (for example, because individuals in the comparison group opted to forgo treatment).

²⁸ In a meta-analysis of juvenile delinquency evaluations, random assignment studies produced effect sizes only 56 percent as large as nonrandom assignment studies. Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy of Political and Social Science*, 587(1), 69-81.

On the 0-to-5 scale as interpreted by the Institute, each study is rated as follows.

- A “5” is assigned to an evaluation with well-implemented random assignment of subjects to a treatment group and a control group that does not receive the treatment/program. A good random assignment study should also indicate how well the random assignment actually occurred by reporting values for pre-existing characteristics for the treatment and control groups.
- A “4” rating is used to designate an experimental random assignment design that had problems in implementation. For example, there could be some crossover between the treatment and control groups or differential attrition rates (such as 10 percent study dropouts among participants versus 25 percent among non-participants).
- A “3” is assigned to an observational study that employs a rigorous quasi-experimental research design with a program and matched comparison group, controlling with statistical methods for self-selection bias that might otherwise influence outcomes. These quasi-experimental methods may include estimates made with a convincing instrumental variables modeling approach, or a Heckman approach to modeling self-selection.²⁹
- A “2” indicates a non-experimental evaluation where the program and comparison groups were reasonably well matched on pre-existing differences in key variables. There must be evidence presented in the evaluation that indicates few, if any, significant differences were observed in these salient pre-existing variables. Alternatively, if an evaluation employs sound multivariate statistical techniques (e.g., logistic regression) to control for pre-existing differences, and if the analysis is successfully completed, then a level “2” study with some differences in pre-existing variables can qualify as a level 3.
- A “1” is used when a level “3” or a “2” study design was less well implemented or didn’t use many statistical controls.
- A “0” involves a study with program and comparison groups that lack comparability on pre-existing variables and no attempt was made to control for these differences in the study. A zero rating also is used in studies where no comparison group is utilized. Instead, the relationship between a program and an outcome, i.e., drug use, is analyzed before and after the program.

We do not use the results from program evaluations rated as a “0” on this scale, because they do not include a comparison group and, thus, no context to judge program effectiveness. In this study, we only considered evaluations that rated at least a 1 on this scale.

B3.2 Adjusting Effect Sizes for Research Involvement in the Program’s Design and Implementation and for Laboratory or Unusual Settings

The purpose of the Institute’s work is to identify and evaluate programs that can make cost-beneficial improvements to Washington’s actual service delivery system. There is some evidence that programs closely controlled by researchers or program developers have better results than those that operate in “real world” administrative structures.³⁰ In our evaluation of a real-world implementation of a research-based juvenile justice program in Washington, we found that the actual results were considerably lower than the results obtained when the intervention was conducted by the originators of the program.³¹ Therefore, we make an adjustment to effect sizes, ES_m , to reflect this distinction. As a parameter for all studies deemed not to be “real world” trials, the Institute discounts ES_m by .5, although this can be modified on a study-by-study basis. We included the “not real world” flag in our regression analyses used to set custom discount rates and used the same procedures to modify this particular discount.

B3.3 Adjusting Effect Sizes for Evaluations With Weak Outcome Measures

Some evaluations use outcome measures that may not be precise gauges of the ultimate outcome of interest. In these cases, we record a flag that can later be used to discount the effect. For example, the evaluation of the Cash and Counseling program³² used a non-standardized survey of clients and caregivers to measure unmet needs, general health, and life satisfaction. If these survey results are used to indicate quality of life, then a flag on this outcome measure can

²⁹ For a discussion of these methods, see Rhodes, W., Pelissier, B., Gaes, G., Saylor, W., Camp, S., & Wallace, S. (2001). Alternative solutions to the problem of selection bias in an analysis of federal residential drug treatment programs. *Evaluation Review*, 25(3), 331-369.

³⁰ Ibid. Lipsey found that, for juvenile delinquency evaluations, programs in routine practice (i.e., “real world” programs) produced effect sizes only 61 percent as large as research/demonstration projects. See also: Petrosino, A., & Soydan, H. (2005). The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1(4), 435-450.

³¹ Barnoski, R. (2004, January). *Outcome evaluation of Washington State’s research-based programs for juvenile offenders* (Document No. 04-01-1201). Olympia: Washington State Institute for Public Policy.

³² Carlson, C. L., Foster, L., & Dale, S. B. (2007). Effects of cash and counseling on personal care and well-being. *Health Services Research*, 42(1, part 2), 467-487.

be used to reflect the probability that this may not be the best measure of quality of life. The same survey was provided to thousands of clients and providers across several states in the Cash and Counseling evaluation, so we included it in our analysis; it would have been better, however, to use a survey that had been standardized before the study was conducted.

B3.4 Values of Adjustment Factors

An explicit adjustment factor (discount rate) is assigned to the results of individual effect sizes based on the Institute’s judgment concerning research quality (study design), research involvement in program design and implementation, “laboratory” setting, and weak outcome measure. Adjustments are made by multiplying the effect size for any study, ES'_m in equation (B5) by the adjustment factors for the topic area. The resulted adjusted effect size is used in the benefit-cost analysis.

For top areas with a limited number of studies, we used the default discount rates listed in Exhibit B1. The default discount rates are subjective to a degree; they are based on the Institute’s general impressions of the confidence that can be placed in the predictive power of evaluations of different quality, weak outcome measures, program developer involvement in evaluation, and unusual settings.

When we had sufficient number of studies in a topic area, we determined discount rates based on results of meta-regression techniques (multivariate linear regress analysis, weighting with random effects inverse variance weights). In many—but not all—topic areas, the discount factors generated by the regression analysis were similar to the default values. For example, in some topic areas, we found no effect of study design on effect size. The detailed tables in Technical Appendix 1 describe the discounts used for each program.

The effect of the discount rates frequently produces a smaller adjusted effect size. For example, using the default adjustments, if a study with ES'_m of -0.20 is deemed a level 4 study, then the -.020 effect size would be multiplied by 0.75 to produce a -0.15 adjusted effect size for use in the benefit-cost analysis.

Exhibit B1
Discount Rates Applied to the Meta-Analysis

| Type of Discount | Discount Rate |
|---|---------------|
| Study Design | |
| 1- Less well-implemented comparison group or observational study, with some covariates. | 0.5 |
| 2- Well-implemented comparison group design, often with many statistical controls. | 0.5 |
| 3- Well-done observational study with many statistical controls (e.g., IV, regression discontinuity). | 0.75 |
| 4- Random assignment, with some RA implementation issues. | 0.75 |
| 5- Well-done random assignment study. | 1.00 |
| Program developer = researcher | 0.5 |
| Unusual (not “real world”) setting | 0.5 |
| Weak measurement used | 0.5 |

Appendix C: Procedures to Compute “Monetizable” Outcome Units

Appendix B described the procedures the Institute uses to compute effect sizes and standard errors. Appendix C describes our procedures to convert effect sizes into units of outcomes that can be monetized. Appendix D then describes how monetary values are attached to these “monetizable” units of outcomes.

To estimate the change in the number of monetizable units for a program or policy, the Institute’s approach draws on two bodies of research: 1) effect sizes from program evaluation research that measure how a program influences an outcome, and 2) effect sizes from other research that estimate the statistical linkage between two different outcomes. The goal is to combine the best current information from these two bodies of research to derive benefit-cost estimates for program and policy choices.

- **Direct Program Effect Sizes for Specific Measured Outcomes.** The first type of effect size measures the estimated direct effect of a program or policy on a particular outcome. We take these direct effect sizes from the original research study itself or, more typically, from a meta analysis of a set of program evaluations on a particular topic. An example of the first type of effect size is an evaluation or meta analysis that directly measures a credible causal relationship between a program such as Nurse Family Partnership and the rate of substantiated child abuse and neglect. In the procedures described below, direct program effect sizes are denoted as PES_o and represent the estimated program effect size on some measured outcome o . The standard error of this effect size, also computed from the original program evaluation or meta analysis, is denoted as $PESSE_o$. Some of these program effect sizes can be monetized directly. To continue the example, a change in substantiated child and abuse can be expected to cause changes in child welfare system costs and in the victimization costs to the child (as described in Appendix D).
- **Linked Effect Sizes on “Monetizable” Outcomes.** The second type of effect size used in the benefit-cost model takes advantage of a different body of research that measures how one particular outcome is causally related to another outcome to which a monetary value can be estimated. An example of the second type of effect size is a (separately estimated) causal linkage between child abuse and neglect and the probability of graduating from high school. Graduating from high school, as described in Appendix D, is an outcome for which monetary benefits can be attached. Thus, while the program itself may have only directly measured an outcome such as child abuse and neglect, the separately estimated linked relationship between child abuse and neglect and high school graduation can be used, in conjunction with the primary research finding, to estimate monetary benefits of the program’s *indirect* effect on high school graduation. The word “indirect” here just means that while the original program evaluation may not have measured a relevant outcome directly, there may be a separate body of credible research indicating a causal relationship between a directly measure outcome and another outcome that can be monetized. In the models below, the “linked” effect sizes are denoted as LES_{om} and represent the estimated effect size between a measured outcome o and a monetizable outcome m . The standard error for this linkage, which is computed from the body of research, is denoted as $LESSE_{om}$.

The procedures outlined below describe how these two types of effect sizes are combined to produce estimates of the units to which monetary values are attached. The product of the procedures is a variable, $Units_m$, that measures the mean number of units of an outcome, m , that can be monetized with the Institute’s benefit-cost model. For example, the units of high school graduation, $Units_{hsgrad}$, might be +0.03, which would indicate three extra percentage points on a high school graduation rate.

C1. Input Screen for Program Effect Size Parameters

The procedures described below use a number of user-supplied parameters.

Some program evaluations measure outcomes just once. For example, a person might be a participant in a program at a certain age and an evaluation measures an outcome at a second age. Some evaluations then measure the same outcome at a subsequent age. Information on how effect sizes change over time can be useful when estimating benefits. The Institute’s benefit-cost model is structured to model an outcome measured at two different post-treatment time intervals. This provides the capability to model program effects that decay, or grow, over time.

The estimated effect of a policy or program on an outcome over time, and the standard error in this estimate, is operationalized in the Institute’s benefit-cost model with eight parameters for each program or policy.

| | |
|--------------|--|
| <i>Tage</i> | the average age of a person treated with a program |
| <i>Mage1</i> | the average age of a person when an effect size of the program is first measured |
| <i>PES1</i> | the estimated effect size of a program at <i>Mage1</i> |

PESSE1 the estimated standard error of the effect size of a program at *Mage1*
Mage2 the average age of a person when an effect size of the program is measured or estimated a second time
PES2 the estimated effect size of a program at *Mage2*
PESSE2 the estimated standard error of the effect size of a program at *Mage2*
BaseRate the estimated outcome for the non-treatment group (e.g., the predicted outcome in absence of the program). For dichotomous outcomes, this is a percentage, for continuous outcomes, it is the standard deviation of the outcome being measured.

Exhibit C.1.a displays a screen shot displaying where these eight parameters are entered for each program, for each outcome. The example shown is the juvenile justice program Functional Family Therapy. The assumed treatment age

Exhibit C1.a

The screenshot shows the 'Enter Program Inputs' tab of the WSIPP Benefit-Cost Model. The program selected is 'FFT (competent) probation', with a short name of 'Functional Family Therapy (Probation)'. The primary participant age is set to 15. The program cost per participant is \$3134 per year. The program description is: 'Functional Family Therapy (FFT) is a structured family-based intervention that uses a multi-step approach to enhance protective factors and reduce risk factors in the family. Functional Family'. The program outcome information table is as follows:

| Outcome | First Effect Size Measurement | | | Second Effect Size Measurement | | | Primary (P) or Secondary (S) | Number of studies in ES estimate | Unadjusted ES at first measurement | Total N in treatment groups | P-value for ES at first measurement |
|---------|-------------------------------|-------------------|-------------------------|--------------------------------|-------------------|--------------------------|------------------------------|----------------------------------|------------------------------------|-----------------------------|-------------------------------------|
| | Effect Size (ES) | ES Standard Error | Age at time of first ES | ES at second measurement | ES Standard Error | Age at time of second ES | | | | | |
| Crime | -0.323 | 0.146 | 16 | -0.323 | 0.292 | 26 | P | 8 | -0.585 | 681 | 0 |

for a juvenile in this program is 15. In the program outcome section of the screen, the user has entered seven of the eight parameters for the crime outcome measured for FFT. The first effect size is -.323 and is measured at age 16 and has a standard error of .146. For this program, our review of the FFT evaluations indicated that the average follow-up period was about one year; thus, we entered age 16 as *Mage1*. The second effect size, -.323, is entered for age 26 with a standard error of .292. The Institute's practice for all programs, such as FFT, that measure an effect size at one follow-up period is to use that adjusted effect size for both the effect size for *Mage1* and *Mage2*, but to double the standard error at *Mage2* to account for the greater uncertainty in the years beyond those measured in the program evaluations. We also set the *Mage2* age ten years beyond the first measured effect size. The eighth parameter, *BaseRate*, is selected by the user by selecting the appropriate population from the drop-down menus in the screen. The actual base rates are entered on other input screens in the software application.

C2. Unit Changes from Direct Effect Sizes

Once these eight parameters are exogenously computed and entered into the model, we follow these steps to compute monetizable units. First, we compute unit changes for outcomes directly measured by the program evaluations.

For dichotomous outcomes:

1. At *Age1* and *Age2*, using the D-cox effect size formula (see Appendix B), we apply *PES1* and *PES2* to the base rates at those two ages to compute the change in monetizable units ($Units_m$) at *Age1* and *Age2*.
2. We then calculate the relative risk ($Units_m / BaseRate$) at *Age1* and *Age2*.
3. For ages ranging from *Age* to *Age1*, we apply the relative risk calculated at *Age1* to the base rates between *Age* and *Age1* to compute the $Units_m$ between *Age* and *Age1*.
4. For ages beyond *Age2*, we apply the relative risk calculated at *Age2* to the base rates after *Age2* to compute the $Units_m$ for all years after *Age2*.
5. For ages ranging from *Age1* to *Age2*, we linearly interpolate the relative risk at each age, and apply that value to the base rates for those ages, to compute the $Units_m$ between *Age1* and *Age2*.

For continuous outcomes:

The unit change ($Units_m$) at each age is simply the effect size (standardized mean difference, see Appendix B) multiplied by the standard deviation unit in which the outcome is measured.

C3. Unit Changes From Linked Effect Sizes

For linked effect sizes, we allow the user to enter a single effect size, standard error, and age of measurement. The unit changes for linked effect sizes are computed as described in the previous section. However, since there is no *Age2*, for dichotomous outcomes, we compute the relative risk ($Units_m / base\ rate$), using the D-cox effect size formula, at *Age1*. We then apply that relative risk to the base rates at all ages (*Age* and beyond). For continuous outcomes, the unit change at each age is simply the effect size at *Age1*, multiplied by the standard deviation unit in which the outcome is measured.

C4. Monetizable Units for Benefit-Cost Calculation

The units of change for effect sizes monetized in the benefit-cost model are simply the multiplicative product of the directly measured (program) and indirect (linked) effect sizes. That is, for a program outcome such as academic test scores, for which do not assume a linkage to other outcomes, the units of change for a program effect size will be the units of change in test scores multiplied by 1. For an outcome such as juvenile crime, for which we estimate a linkage to high school graduation, we calculate two sets of unit changes. For the direct (crime) measure, we simply use the unit change for crime multiplied by 1. For the indirect (high school graduation) measure, we multiply that unit change in crime by the unit change for the link between crime and high school graduation.

Appendix D: Methods Used to Estimate Monetary Benefits

As summarized in Appendix A, the Institute model is an integrated set of estimates and computational routines designed to produce internally consistent benefit-to-cost estimates for a variety of public policies and programs. The model implements a standard economic calculation of the expected worth of an investment by computing the net present value of a stream of estimated benefits and costs that occur over time, as described with equation (D1).

$$(D1) \quad NPV_{tage} = \sum_{y=tage}^N \frac{Q_y \times P_y - C_y}{(1 + Dis)^y}$$

In this basic model, the net present value (*NPV*) of a program is the quantity of the outcomes produced by the program or policy (*Q*) in year *y*, times the price per unit of the outcome (*P*) in year *y*, minus the cost of producing the outcome (*C*) in year *y*. The lifecycle of each of these values is present-valued to the average age a person is treated (*tage*) and covers the number of years into the future over which they are evaluated (*N*). The future values are expressed in present value terms after applying a discount rate (*Dis*). An internal rate of return on investment can also be calculated from these annual cash flows. As noted, many of the values summarized in equation (D1) are estimated or posited with uncertainty; we model this uncertainty using a Monte Carlo simulation to estimate the riskiness of benefit-cost results.

The first term in the numerator of equation (D1), Q_y , is the estimated number of outcome “units” in year *y* produced by the program or policy. The procedures we use to develop estimates of Q_y are described in Appendices B and C. In Appendix D we describe the various methods we use to estimate the price term, P_y , in equation (D1).

D1. Valuation of Outcomes That Affect Labor Market Earnings

Several of the outcomes measured in the benefit-cost model are monetized, in part, with labor market earnings. Measuring the earnings implications of human capital variables is a common approach in economics.³³

In the current version of the benefit-cost model, the following outcomes are monetized, in part, with labor market earnings:

- High school graduation
- Standardized student test scores
- Number of years of completed education
- Morbidity and mortality costs of alcohol and illicit drug disorders, and regular smoking
- Morbidity and mortality costs of mental health disorders

This section discusses the data sources we use for estimates of labor market earnings. Other parts of Appendix D present additional parameters for the specific outcomes listed above, along with the computational routines to produce labor market earnings benefits.

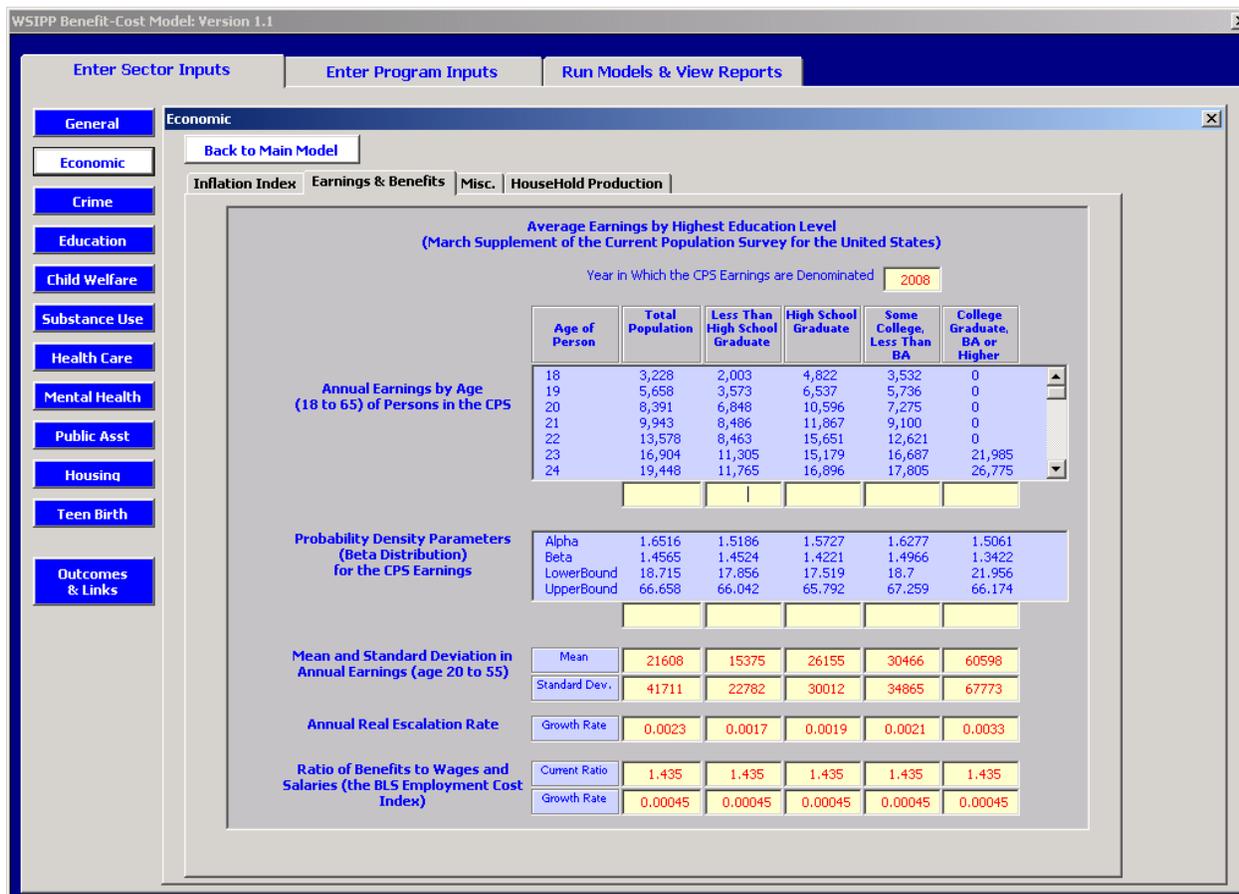
D1.1 Earnings Data and Related Parameters

In this analysis, all earnings estimates derive from a common dataset. The estimates are taken from the U.S. Census Bureau’s March Supplement to the Current Population Survey (CPS), which provides, annually, cross sectional data for earnings by age and by educational status.³⁴ These data are representative of the United States population, not just those living in Washington State. Exhibit D1.a shows an input screen from the Institute’s cost-benefit model that displays the CPS data and related parameters used in the benefit-cost model.

³³ See, for example, Heckman et al., 2010. See also, Rouse, C. E. (2007). Consequences for the labor market. In C. R. Belfield & H. M. Levin (Eds.), *The price we pay: Economic and social consequences of inadequate education* (pp. 99-124). Washington, DC: Brookings Institution.; Krueger, A. B. (2003). Economic considerations and class size. *The Economic Journal*, 113(485), F34-F63; and Hanushek, E. A. (2004, January). *Some simple analytics of school quality* (NBER Working Paper No. 10229). Cambridge, MA: National Bureau of Economic Research.

³⁴ The data are accessed from the “DataFerrett” application of the U. S. Department of Commerce, Bureau of the Census, available from <http://dataferrett.census.gov>.

Exhibit D1.a



From a recent annual March supplement to the CPS, we collect average personal earnings by age of each person and by educational status. We gather the following three “Person Variables” from the CPS: (1) PEARINVAL, Person Total earnings—this variable measures income from earnings, not total money income; (2) A_AGE, age by single year; and (3) A_HGA, educational attainment by the highest level completed. From these data we compute average earnings per person, by single year of age for five educational status groupings:

- the total population—that is, the entire CPS sample (Institute variable name: *CPSEarnAll*);
- those who did not report completing high school but completed 7th grade or higher (*CPSEarnNHSG*);
- those who reported completing high school with a diploma (*CPSEarnHSG*);
- those with some college, but not a BA degree (*CPSEarnSomeCol*); and
- those with a BA degree or higher (*CPSEarnBA+*).

It is important to note that the average earnings reported are for all people at each age, not just for those with earnings. Thus, the data series measure both earnings of the earners and the rate of labor force participation.

From these five annual earnings streams for a recent CPS year (for example, the 2009 CPS report contains data for 2008 earnings), we fit probability density distributions. We use Palisade Corporation’s *@Risk* program to select the probability distribution with the lowest root mean square error. For all five series, we found the best probability distribution to be a beta distribution. The four fitted beta distribution parameters (Alpha, Beta, LowerAge, and UpperAge) for these distributions are then entered into the model, as shown in Exhibit D1.a. These beta distributions are used to allocate the sum of all cross-sectional total earnings reported for all ages for the particular education cohort. For example, for the annual earnings estimates for the total population in the CPS sample (*CPSEarnAll*), we compute the following for each year *y*:

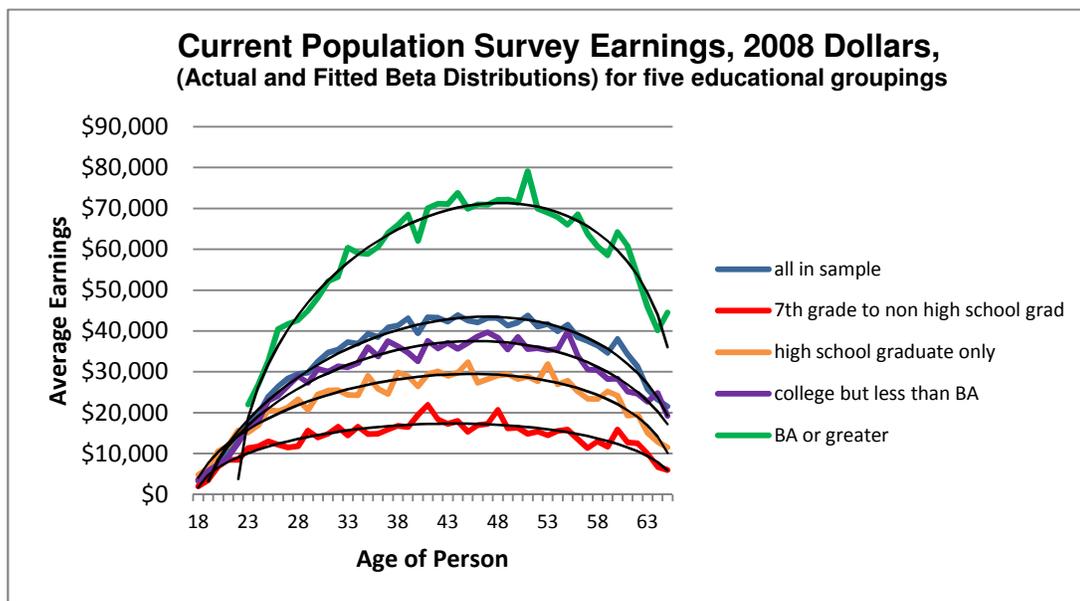
$$EarnAll_y = \left[\sum_{t=18}^{65} CPSEarnAll_t \right] \times \left[\frac{(y - LowerAge)^{ALPHA-1} \times (UpperAge - y)^{BETA-1}}{B(ALPHA, BETA) \times (UpperAge - LowerAge)^{ALPHA+BETA-1}} \right]$$

Where *ALPHA* and *BETA* are the estimated shape parameters for the beta distribution for the total population CPS earnings, and *LowerAge* and *UpperAge* are the estimated continuous bounding parameters for the total population CPS earnings. *B* is the beta function which can be calculated in Microsoft Excel for the total population CPS earnings with:

$$B(alpha, beta) = \frac{EXP[GAMMALN(alpha)] \times EXP[GAMMALN(beta)]}{EXP[GAMMALN(alpha + beta)]}$$

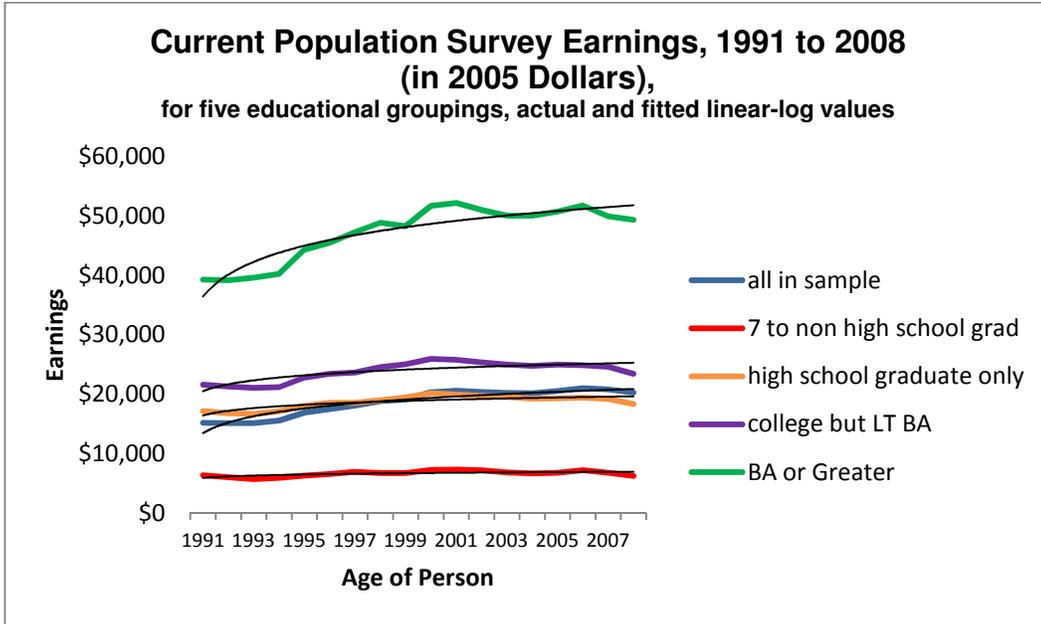
The same process is used to estimate the annual CPS earnings streams for the four other educational achievement groups, substituting the relevant parameters for each group. The raw CPS earnings data, along with the fitted curves from these procedures are plotted below.

Exhibit D1.b



Growth Rates in Earnings. Since these CPS data are cross sections for the most recent CPS year, and since our benefit cost analysis reflects life-cycle earnings, we also compute an estimate of the long-run real rate of escalation in earnings for each of the five groups. We collect the same cross-sectional CPS information for the five groups for all of the years electronically available from the Census website: 1992 (with data for 1991) to 2009 (with data for 2008). We adjust each series for inflation using the United States Implicit Price Deflator for Personal Consumption Expenditures from the U.S. Department of Commerce, as published and forecasted by the Washington Economic and Revenue Forecast Council. We then fit linear-log models (earnings = a + b(ln(year)) to each of the five series. The actual data and the fitted linear-log models are shown on the following chart.

Exhibit D1.c



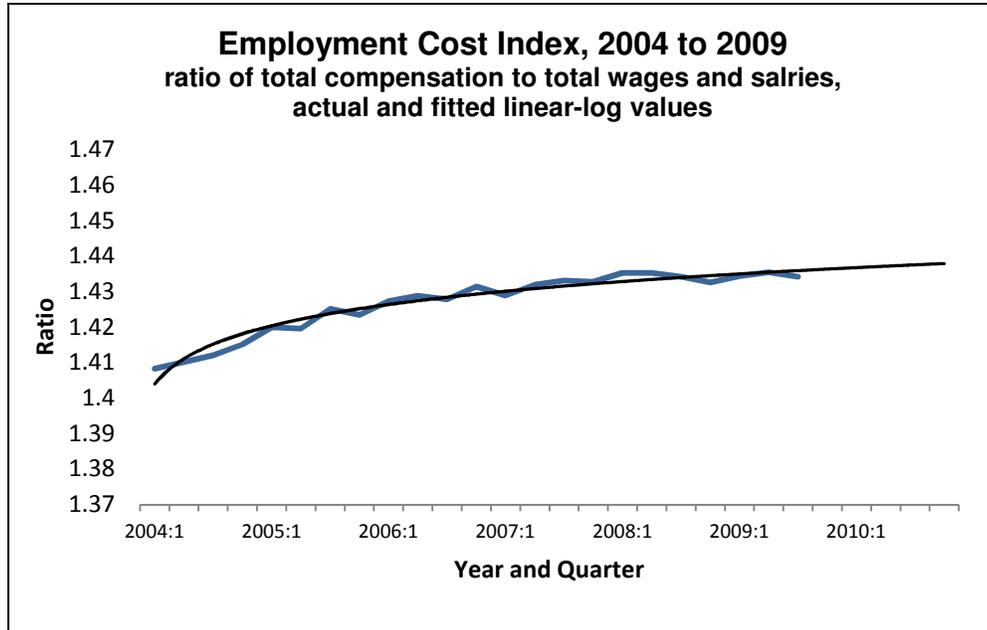
From the linear-log regression coefficients, we then estimate the values for 2008 and 2038, and compute a forecast of the annual rate of real growth in earnings over the 30-year forecasted interval. These estimates of the annual real rate of change in wages are then entered into the model, as shown on D1.a.

Employee Benefits. The CPS data are for earnings and do not include employee benefits associated with earnings. To measure these additions to earnings, we include an estimate of the ratio of total employee compensation to wage and salaries. We compute these estimates from the Bureau of Labor Statistics (BLS) Employer Cost Index (ECI).³⁵ According to the Bureau of Labor Statistics, the benefits covered by the ECI are: “Paid leave--vacations, holidays, sick leave, and personal leave; supplemental pay--premium pay for work in addition to the regular work schedule (such as overtime, weekends, and holidays), shift differentials, and nonproduction bonuses (such as year-end, referral, and attendance bonuses); insurance benefits--life, health, short-term disability, and long-term disability; retirement and savings benefits--defined benefit and defined contribution plans; and legally required benefits--Social Security, Medicare, federal and state unemployment insurance, and workers’ compensation.”

The chart below displays the quarterly national ECI ratio of total compensation to total wages for all civilian workers. We fit a linear-log model ($\text{ratio} = a + b(\ln(\text{quarter}))$) to the series and estimate the annual values for 2008 and 2038, and then compute a forecast of the annual rate growth in the benefit ratio over the 30 year interval. The current year value and the growth rate are then entered into the model, as shown on D1.a. Unfortunately, the current BLS ECI does not allow the index or the growth rate to be broken out by education achievement level. Therefore, the same values are entered for each group. It is plausible that there are differences in the base rate and the expected growth rate in benefits by educational level. The model is structured so that these parameters can be included in the future.

³⁵ U. S. Bureau of Labor Statistics. (2008, June 11). *Employer costs for employee compensation—March 2011* (USD11-0849), Washington, DC: Author. Retrieved June 30, 2011 from <http://www.bls.gov/news.release/ecec.toc.htm>

Exhibit D1.d



The earnings series is then used in the benefit-cost model to estimate labor market-related benefits of a number of outcomes. For example, in each year (y), the basic CPS earnings series is adjusted with the factors:

$$ModEarnAll_y = (EarnAll_y \times (1 + EscAll)^{y-tage}) \times (Fall \times (1 + EscFall)^{y-tage}) \times (IPD_{base} / IPD_{cps})$$

In this example, for each year (y) from the age of a program participant ($tage$) to age 65, the annual CPS earnings for all people ($EarnAll$) are multiplied by one plus the relevant real earnings escalation rate ($EscAll$) raised to the number of years after program participation, times the fringe benefit rate for all people ($Fall$), multiplied by one plus the relevant fringe benefit escalation rate ($EscFall$) raised to the number of years after program participation, times a factor to apply the Implicit Price Deflator for the base year dollars (IPD_{base}) chosen for the overall benefit-cost analysis relative to the year in which the CPS data are denominated (IPD_{cps}).

D2. Valuation of Outcomes That Affect Crime

This section of the Appendix describes the Institute's benefit-cost model that estimates the monetary value to taxpayers and victims of programs that reduce crime. In this Appendix, we describe the methods, data sources, and estimation procedures.

The current version of the Institute's model approaches the crime valuation question from two perspectives. We compute the value to taxpayers if a crime is avoided. We also estimate the costs that can be avoided by people who would otherwise have been a victim of a crime, had the crime not been averted.³⁶ To model avoided crime costs from these two perspectives, we estimate life-cycle costs of avoiding seven major types of crime and 11 types of costs incurred as a result of crime. In addition to computing monetary values of avoided crime, the model is also used to estimate and count the number of prison beds and victimizations avoided when crime is reduced.

We also developed a "sentencing and corrections" module to help Washington, and perhaps other states, identify evidence-based policy mixes that can both fight crime and save taxpayers money; however we do not describe the sentencing module in this appendix.³⁷ The sentencing module, which resides within the Institute's larger benefit-cost model, helps users analyze a portfolio of criminal justice sentencing and programming policies.

The crime model uses four broad types of inputs: per-unit crime costs; sentencing probabilities and resource-use estimates; longitudinal criminological information about different populations; and estimates of multiple crimes per officially recorded crimes, such as arrests or convictions. This section begins by describing these four data sources and then turns to the computational procedures that produce the avoided costs of reduced crime.

D2.1 Per-Unit Crime Costs

In the Institute's benefit-cost model, the costs of the criminal justice system paid by taxpayers are estimated for each significant part of the publicly financed system in Washington. The sectors modeled include the costs of police and sheriffs, superior courts and county prosecutors, local juvenile corrections, local adult corrections, state juvenile corrections, and state adult corrections. The estimated costs include operating costs and annualized capital costs for the capital-intensive sectors. As noted, we also include estimates of the costs of crime to victims.

For criminal justice system costs, the estimates are *marginal* operating and capital costs.³⁸ Marginal criminal justice costs are defined as those costs that change over a period of several years as a result of changes in a crime workload measure. Some short-run costs change instantly when a workload changes. For example, when one prisoner is added to the state adult corrections' system, certain variable food and service costs increase immediately, but new staff are not typically hired right away. Over the course of a governmental budget cycle, however, new corrections' staff are likely to be hired to reflect the change in average daily population of the prison. In the Institute's analysis, these "longer-run" marginal costs have been estimated. The longer-run marginal costs reflect both the immediate short-run changes in expenditures, as well as those operating expenditures that change after governments make adjustments to staffing levels, often in the next few budget-writing cycles.

³⁶ There are other costs of crime that have been posited by some commentators and analysts, including private costs and other public sector costs. WSIPP's current model does not address these additional cost categories or does so only indirectly. Future versions of this model may incorporate some of these additional cost categories.

³⁷ See: Aos, S., & Drake, E. (2010, August). *WSIPP's benefit-cost tool for states: Examining policy options in sentencing and corrections* (Document No. 10-08-1201). Olympia: Washington State Institute for Public Policy.

³⁸ As noted, a few average cost figures are currently used in the model when marginal cost estimates cannot be reasonably estimated.

Exhibit D2.a shows a screen shot, taken from the Institute's benefit-cost model, that displays an array of per-unit costs for the 12 sectors and seven types of crime modeled. The estimates for each row in Exhibit D2.a are described below.

Exhibit D2.a

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

General | Economic | **Crime** | Education | Child Welfare | Substance Use | Health Care | Mental Health | Public Asst | Housing | Teen Birth | Outcomes & Links

Crime

Back to Main Model

Per Unit Costs | Resource Use | Offender Populations | Victimization | Program Participation

| | Marginal Operating Costs | | | | | | | Capital Costs | | | Misc. | | |
|-------------------------------|--------------------------|-------------------|---------|--------------------|-----------------|-------------|-------------|------------------|----------------------|-----------------------|------------------|---------------|-----------------------|
| | Murder | Felony Sex Crimes | Robbery | Aggravated Assault | Felony Property | Felony Drug | Misdemeanor | Year of Estimate | Real Escalation Rate | Capital Cost Per Unit | Year of Estimate | Finance Years | Percent Paid by State |
| Police | 670 | 670 | 670 | 670 | 670 | 670 | 670 | 2009 | 0.0270 | 0 | 2006 | 5 | 0% |
| Courts and Prosecutors | 152,378 | 18,770 | 9,865 | 4,877 | 201 | 201 | 201 | 2009 | 0.0200 | 370 | 2006 | 20 | 0% |
| Juvenile Local Detention | 20,293 | 20,293 | 20,293 | 20,293 | 20,293 | 20,293 | 20,293 | 2009 | 0.0570 | 200,000 | 2009 | 25 | 0% |
| Juvenile Local Supervision | 5,200 | 5,200 | 5,200 | 5,200 | 5,200 | 5,200 | 5,200 | 2008 | 0.0000 | | | | 0% |
| Juvenile State Institution | 36,743 | 36,743 | 36,743 | 36,743 | 36,743 | 36,743 | 36,743 | 2009 | 0.0160 | 150,000 | 2009 | 25 | 100% |
| Juvenile State Supervision | 3,927 | 3,927 | 3,927 | 3,927 | 3,927 | 3,927 | 3,927 | 2009 | 0.0000 | | | | 100% |
| Adult Jail | 21,469 | 21,469 | 21,469 | 21,469 | 21,469 | 21,469 | 21,469 | 2009 | 0.0220 | 150,000 | 2009 | 25 | 0% |
| Adult Local Supervision | 1,861 | 1,861 | 1,861 | 1,861 | 1,861 | 1,861 | 1,861 | 2009 | 0.0640 | | | | 100% |
| Adult State Prison | 12,722 | 12,722 | 12,722 | 12,722 | 12,722 | 12,722 | 12,722 | 2009 | 0.0030 | 113,339 | 2007 | 25 | 100% |
| Adult Post Prison Supervision | 1,861 | 1,861 | 1,861 | 1,861 | 1,861 | 1,861 | 1,861 | 2009 | 0.0640 | | | | 100% |
| Victim Costs (tangible) | 737,517 | 5,556 | 3,299 | 8,700 | 1,922 | 0 | 0 | 2008 | 0.0000 | | | | 0% |
| Victim Costs (intangible) | 8,442,000 | 198,212 | 4,976 | 13,435 | 0 | 0 | 0 | 2008 | 0.0000 | | | | 0% |

Notes: Police costs are dollars per arrest. Courts costs (including court, prosecutors, and defenders) are dollars per conviction. Victim costs are present value costs per victim. All other costs are annual costs per average daily population unit.

Cost Variance for Per Unit Criminal Justice and Victim Costs

| | Low Variance | High Variance |
|-------------------------------|--------------|---------------|
| Criminal Justice System Costs | -0.1 | 0.1 |
| Crime Victim Costs | -0.2 | 0.1 |

Police and Sheriff's Office Per-Unit Costs

This section describes the steps we use to estimate the annual marginal operating costs of local police agencies in Washington State, along with the expected long-run real rate of change in these costs. We also describe our estimate of the capital cost of police operations. All of these cost parameters are entered into the crime model, as shown in Exhibit D2.a.

Police Operating Costs. For an estimate of marginal operating costs of local police agencies, we conducted a time-series analysis of annual county-level data for police expenditures and arrests for all local police agencies in Washington's 39 counties. From the Washington State Auditor, local city and county police expenditure data were collected for 1994 to 2008, the earliest and latest years, as of winter 2010, electronically available. The Auditor's data for the expenses include all local police expenditures (Budget and Reporting System (BARS) code 521). We excluded the Crime Prevention (BARS 521.30) subcategory since it was an irregular expenditure. These nominal annual dollar amounts were adjusted to 2009 dollars using the United States Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce, as published and forecasted by the Washington Economic and Revenue Forecast Council.

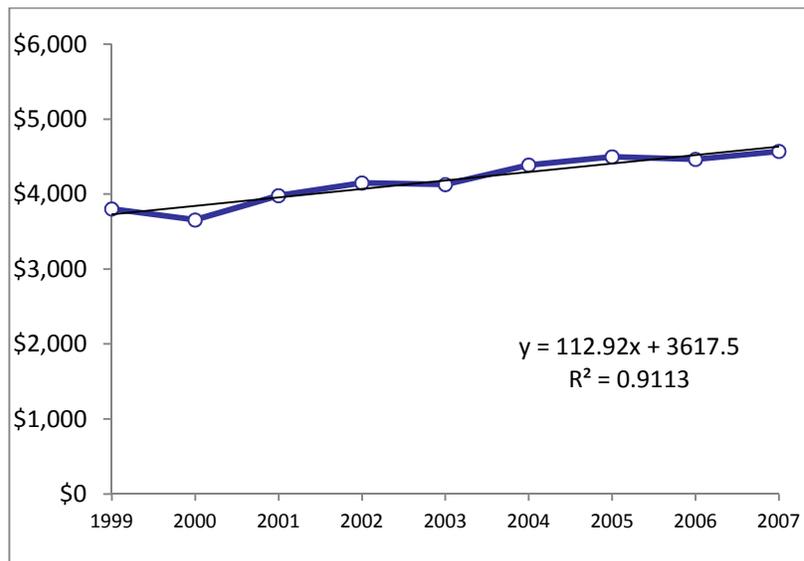
We also collected arrest information for Washington police agencies from the National Archive of Criminal Justice Data maintained by the University of Michigan.³⁹ Data were collected for calendar years 1994 to 2007, the earliest and latest years available as of December 2009. Arrest data for 1993 were unavailable on the Michigan website, thus limiting the number of years we could include in our analysis.

We aggregated the city and county expenditure and arrest data for individual police agencies to the county level to account for any jurisdictional overlap in county sheriffs' offices and city police units. We also aggregated to the county level because, over the years included in our analysis, some newly incorporated cities took on responsibilities formerly assigned to county sheriffs. Aggregating thus allowed for a more consistent cost-arrest data series for the years in our study. Since the latest arrest data were for 2007, the resulting balanced multiple time-series panel dataset initially consisted of 546 county-by-year observations.

We had to limit our analysis to 1999 to 2007 because visual inspection of the arrest data for years 1996 to 1998 revealed what appeared to be significant anomalies in the data, possibly due to reporting or other unknown factors during those years. Therefore, in our regression analyses, our dataset begins in 1999.

We computed the statewide average cost per arrest (in 2009 dollars) for 1999 to 2007 and plotted the results.

Exhibit D2.b
Average Police Costs Per Arrest, 2009 Dollars
Calendar Years 1999 to 2007



³⁹United States Department of Justice, Federal Bureau of Investigation. *Uniform crime reporting program data [United States]: County-level detailed arrest and offense data* [by year]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Over the entire 1999 to 2007 timeframe, the average statewide cost is \$4,182 per arrest, in 2009 dollars. Over these years, there has been an upward trend in the inflation-adjusted costs. We computed an estimate of the average annual real escalation rate in costs by estimating a linear line (shown in the chart) for this series. From this line, we computed the predicted values for 1999 (\$3,734) and 2007 (\$4,638) and calculated the average escalation rate for the eight years, using the following formula, where FV is the 2007 estimated cost, PV is the 1999 estimate, and N is eight years.

$$(1) \text{ Rate} = (FV/PV)^{1/N}$$

The annual rate of real escalation is .027. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a.

Next, to estimate the marginal annual operating costs of police agencies, we conducted a time-series analysis of the panel data for Washington's 39 counties for 2001 to 2007. The restriction to 2001 (for the dependent variable) was because, after testing different lag structures of right-hand side variables, and to preclude using arrest data before 1999, our sample dependent variable began in 2001. Thus the balanced panel includes a total of 273 observations (39 counties for 7 years). We tested models where we disaggregated the arrest data into five types: arrests for murder, rape, robbery, aggravated assault, and all nonviolent arrests. After testing a variety of specifications, we did not find a specification with stable or intuitively reasonable results. At this time, we do not know if there are measurement errors in the arrest data, or if there are other tests to be explored. Therefore, we estimated a simple model with total arrests. This model, however, is unsatisfactory because it implies, for example, that the cost for an arrest for murder is the same as the cost for an arrest for burglary. We intend to examine the historical arrest data in greater detail so that a more intuitive equation can be estimated with disaggregated arrest types. The arrest data do not include the traffic operations of local police agencies. To capture this effect, data from the Washington State Administrative Office of the Courts were obtained on the number of traffic infraction filings in county courts.

In our time series analysis, we first tested each data series for unit roots. The data series are: real police expenditures (M_POLICER), total arrests (A_TOT), and traffic infractions (TRAFFIC). If unit roots are present, then a simple regression in levels can produce spurious results.⁴⁰ We tested for unit roots with the Im, Pesaran, and Shin (IPS) panel unit root test for individual unit root processes.

- For the M_POLICER expenditure series, the IPS test without time trends failed to reject the null hypotheses that the series has a unit root (IPS p-value of 1.00). With time trends included, the IPS test continued to indicate a unit root (IPS p-value .34). In first-differences, on the other hand, the IPS test indicated a lack of a unit root (IPS p-value .000).
- For the two right-hand side variables, the IPS tests indicated a lack of a unit root for A_TOT (IPS p-value of .000), but a unit root for TRAFFIC (IPS p-value of .88).
- With the IPS test indicating a unit root in the dependent variable (M_POLICER), we proceeded to construct a model in first-differences.

We tested alternative lag specifications of the arrest and traffic variables. Our preferred model also included period and county fixed effects and a lagged dependent variable. The following results were obtained and the coefficients entered in the crime model, as shown in Exhibit D2.a. The sum of the arrest lags is \$670. An identical model, but without including a right-hand side dependent variable, produced quite similar results.

⁴⁰ Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach*. Mason, OH: South-Western Cengage Learning, p. 636.

Exhibit D2.c

Dependent Variable: M_POLICER-M_POLICER(-1)

Method: Panel Least Squares

Date: 04/17/10 Time: 10:29

Sample (adjusted): 2001 2007

Periods included: 7

Cross-sections included: 39

Total panel (balanced) observations: 273

White period standard errors & covariance (d.f. corrected)

WARNING: estimated coefficient covariance matrix is of reduced rank

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|-----------------------------|-------------|------------|-------------|--------|
| C | 956767.2 | 171084.1 | 5.592380 | 0.0000 |
| M_POLICER(-1)-M_POLICER(-2) | -0.468607 | 0.097310 | -4.815585 | 0.0000 |
| A_TOT-A_TOT(-1) | 240.6135 | 331.7045 | 0.725385 | 0.4690 |
| A_TOT(-1)-A_TOT(-2) | 428.8218 | 319.8050 | 1.340886 | 0.1813 |
| TRAFFIC-TRAFFIC(-1) | 109.2628 | 87.19574 | 1.253075 | 0.2115 |
| TRAFFIC(-1)-TRAFFIC(-2) | 123.4954 | 97.02971 | 1.272759 | 0.2044 |
| TRAFFIC(-2)-TRAFFIC(-3) | 350.3366 | 115.0134 | 3.046049 | 0.0026 |

Effects Specification

Cross-section fixed (dummy variables)

Period fixed (dummy variables)

| | | | |
|--------------------|-----------|-----------------------|----------|
| R-squared | 0.679778 | Mean dependent var | 1013022. |
| Adjusted R-squared | 0.607657 | S.D. dependent var | 3244727. |
| S.E. of regression | 2032410. | Akaike info criterion | 32.05417 |
| Sum squared resid | 9.17E+14 | Schwarz criterion | 32.72847 |
| Log likelihood | -4324.395 | Hannan-Quinn criter. | 32.32485 |
| F-statistic | 9.425402 | Durbin-Watson stat | 1.964607 |
| Prob(F-statistic) | 0.000000 | | |

Police Capital Costs. An estimate of the capital costs used by local police to make arrests in Washington was calculated from capital expenditure data for local police agencies in Washington for 2006. These data were obtained from the United States Bureau of Justice Statistics annual survey: Justice Expenditure and Employment Extracts, 2006, published December 1, 2008 (NCJ 224394). Local government police capital expenditures in Washington were reported as \$53,703,000 for 2006 (Table 4, Justice system expenditure by character, State and type of government, fiscal 2006). The total number of arrests in Washington during 2006 was 246,388, obtained from FBI's Uniform Crime Reports for 2006. Thus, the average police capital cost per arrest was \$218 in 2006 dollars. This parameter was entered into the crime model, as shown in Exhibit D2.a, along with an assumed five-year financing for these police resources. In our crime model, the total capital cost per arrest is converted to an annualized capital payment, with equation (2), assuming a five-year financing term (n), the bond financing rate entered in the model (i), and setting PV equal to the capital cost per arrest converted to the base year dollars chosen for the model.

$$(2) PMT = \frac{iPV}{1 - (1 + i)^{-n}}$$

Superior Courts and County Prosecutors Per-Unit Costs

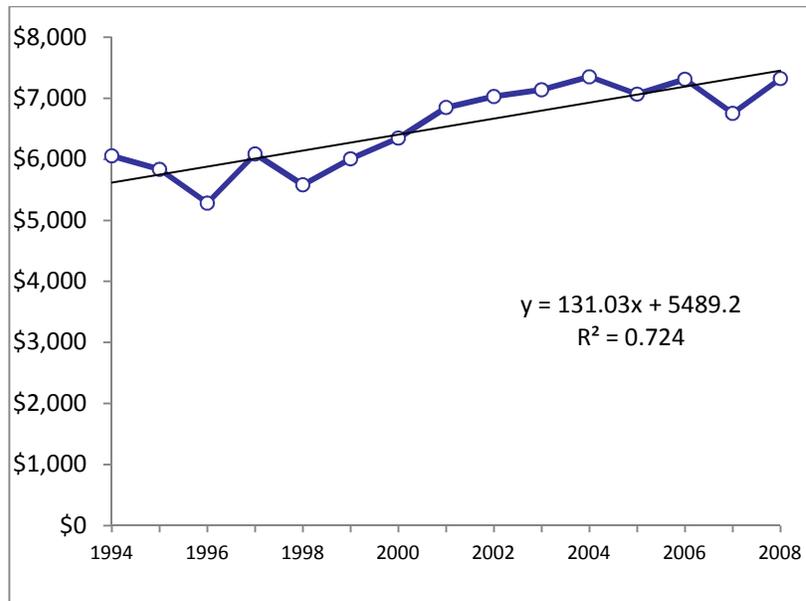
This section describes the steps we use to estimate marginal annual operating costs, and the long-run rate of change in these costs, of county superior courts and prosecutors in Washington State. Our focus is the cost of obtaining convictions in courts, so we combined court costs and prosecutor costs into one category to reflect the public costs to process cases through the courts that respond especially to felony crime. The cost parameters are entered into the crime model, as shown in Exhibit D2.a.

Court and Prosecutor Operating Costs. For an estimate of marginal operating costs of superior courts in Washington, we conducted a time series analysis of annual county-level data for court and prosecutor expenditures and court convictions for all local agencies in Washington's 39 counties. From the Washington State Auditor, local county court and prosecutor expenditure data were collected for calendar years 1994 to 2008, the earliest and latest years, as of winter 2010, electronically available. The Auditor's data for the expenses includes all local court and prosecutor expenditures (BARS code 512 for courts and BARS code 515 for prosecutors). The court data include the costs of administration (BARS 512.10), superior courts (BARS 512.20), and county clerks (BARS 512.30). For court expenditure data, we excluded district courts (BARS 512.40), since they do not process felony cases (the main subject of interest in our benefit-cost analysis) and expenditures for law library (BARS 512.70) and indigent defense (BARS 512.80); this latter category was excluded because the data were not available for the entire time frame under review. The prosecutor data include costs for administration-legal (515.10) and legal services (515.2). For prosecutor offices, we excluded facilities-legal services (515.50), consumer affairs-legal services (515.60), crime victim and witness program-legal (515.70), and child support enforcement-legal services (515.80). All nominal annual dollar amounts were adjusted to 2009 dollars using the United States Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce, as published and forecasted by the Washington Economic and Revenue Forecast Council.

We also collected court conviction and other case-processing information from the Washington State Administrative Office of the Courts. We collected statewide data for calendar years 1994 to 2008 and county-level data for calendar years 1997 to 2008, the earliest and latest years available as of December 2009.

We computed the statewide average cost per conviction (in 2009 dollars) for 1994 to 2008 and plotted the results.

Exhibit D2.d
Average Court Costs Per
Conviction, 2009 Dollars
Calendar Years 1994 to 2008



Over the entire 1994 to 2008 timeframe, the average statewide cost is \$6,557 per conviction, in 2009 dollars. Over these years, there has been an upward trend in the inflation-adjusted costs. We computed an estimate of the average annual real escalation rate in costs by estimating a linear line (shown in the chart) for this series. From this line, we computed the predicted values for 1994 (\$5,625) and 2008 (\$7,461) and calculated the average escalation rate for the 14 years, using equation (1), where FV is the 2008 estimated cost, PV is the 1994 estimate, and N is 14 years.

The annual rate of real escalation is .020. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a.

Next, to estimate the marginal annual operating costs of courts, we conducted a time-series analysis of the panel data for Washington’s 39 counties for 1999 to 2008. The restriction to 1999 (for the dependent variable) was because, after testing different lag structures of right-hand side variables, and since our county-level court data began in 1997, our sample dependent variable had to begin in 1999. Thus, the balanced panel includes a total of 390 observations (39 counties for 10 years). Conviction data were categorized into four types of violent convictions and one for all other convictions.

In our time-series analysis, we first tested each data series for unit roots. The six data series are: real total court expenditures (M_COURTALLR), convictions for homicide offenses (C_HOM), convictions for sex offenses (C_SEX), convictions for robbery offenses (C_ROB), convictions for aggravated assault offenses (C_ASSLT), convictions for all non-violent offenses (C_NONVIOL). If unit roots are present, then a simple regression in levels can produce spurious results.⁴¹ We tested for unit roots with the Im, Pesaran, and Shin (IPS) panel unit root test for individual unit root processes.

- For all of the variables, the IPS tests generally indicated a lack of unit roots. For example, IPS test without time trends rejected the null hypotheses that the series have unit roots (IPS p-values of .0028 for M_COURTALLR, .0000 for C_HOM, .0000 for C_SEX, .0000 for C_ROB, .0000 for C_ASSLT, .0006 for C_NONVIOL).
- With the IPS test indicating a lack of unit roots in the variables, we had the option to construct models in levels or first-differences.

⁴¹ Ibid., p. 636.

We tested models both in levels and first-differences, along with alternative lag specifications for the conviction variables. Our preferred model was a first-difference model where we included lags of each of the violent felony conviction variables along with a variable for all other convictions, as well as county and time fixed effects. We also included a lagged dependent variable. This model produced coefficients for the violent conviction variables that made the most intuitive sense.

Exhibit D2.e

| Dependent Variable: M_COURTALLR-M_COURTALLR(-1) | | | | |
|--|-------------|-----------------------|-------------|--------|
| Method: Panel Least Squares | | | | |
| Date: 02/04/10 Time: 10:01 | | | | |
| Sample (adjusted): 1999 2008 | | | | |
| Periods included: 10 | | | | |
| Cross-sections included: 39 | | | | |
| Total panel (balanced) observations: 390 | | | | |
| White diagonal standard errors & covariance (d.f. corrected) | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 158006.5 | 86235.19 | 1.832274 | 0.0678 |
| M_COURTALLR(-1)-M_COURTALLR(-2) | -0.113178 | 0.168569 | -0.671403 | 0.5024 |
| C_HOM(-1)-C_HOM(-2) | 152377.9 | 125366.9 | 1.215456 | 0.2250 |
| C_SEX(-1)-C_SEX(-2) | 18770.28 | 11395.58 | 1.647154 | 0.1005 |
| C_ROB(-1)-C_ROB(-2) | 9865.480 | 29782.45 | 0.331252 | 0.7407 |
| C_ASSLT(-1)-C_ASSLT(-2) | 4876.710 | 9512.385 | 0.512670 | 0.6085 |
| C_NONVIOL-C_NONVIOL(-1) | 200.5611 | 1503.985 | 0.133353 | 0.8940 |
| Effects Specification | | | | |
| Cross-section fixed (dummy variables) | | | | |
| Period fixed (dummy variables) | | | | |
| R-squared | 0.209477 | Mean dependent var | 167352.1 | |
| Adjusted R-squared | 0.084781 | S.D. dependent var | 2196761. | |
| S.E. of regression | 2101577. | Akaike info criterion | 32.08216 | |
| Sum squared resid | 1.48E+15 | Schwarz criterion | 32.63132 | |
| Log likelihood | -6202.021 | Hannan-Quinn criter. | 32.29985 | |
| F-statistic | 1.679903 | Durbin-Watson stat | 1.973011 | |
| Prob(F-statistic) | 0.003621 | | | |

Court Capital Costs. An estimate of the capital costs used by the court system in Washington was calculated from capital expenditure data for courts in Washington for 2006. These data were obtained from the United States Bureau of Justice Statistics annual survey: Justice Expenditure and Employment Extracts, 2006, published December 1, 2008 (NCJ 224394). Local government court expenditures in Washington were reported as \$19,144,000 for 2006 (Table 4, Justice system expenditure by character, State and type of government, fiscal 2006). The total number of criminal (adult and juvenile) convictions in Washington during 2006 was 51,709, obtained from the Washington State Administrative Office of the Courts. Thus, the average court capital cost per conviction was \$370 in 2006 dollars. This parameter was entered into the crime model, as shown in Exhibit D2.a, along with an assumed 20-year financing period. In our crime model, the total capital cost per conviction is converted to an annualized capital payment, with equation (2), assuming a 20-year financing term (n), the bond financing rate entered in the model (i), and setting PV equal to the capital cost per conviction converted to the base year dollars chosen for the model.

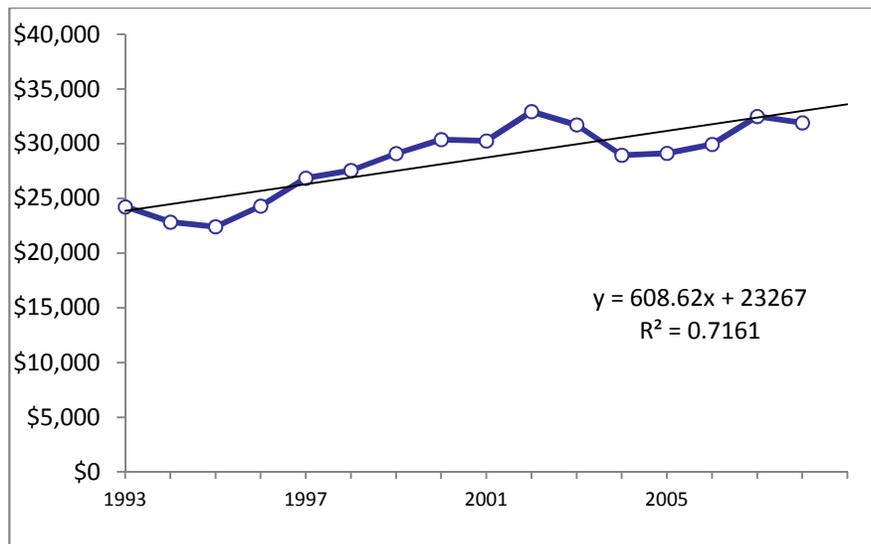
Local Adult Jail Per-Unit Costs

This section describes the steps we use to estimate marginal annual jail operating costs, and the long-run rate of change in these costs, of the county-run adult jail system in Washington State. We also describe our estimate of the capital cost per jail bed. All of these cost parameters are entered into the crime model, as shown in Exhibit D2.a. In the Institute's model, two types of users of local county-run adult jails are analyzed: convicted felons who serve both pre-sentence and post-sentence time at a local jail, and felons who serve pre-sentence time at local jails and post-sentence time at a state institution. The Institute assumes the same annualized per-day jail cost for both these events.

Jail Operating Costs. For an estimate of marginal operating costs of county jails, we conducted a time-series analysis of annual county-level data for jail expenditures and average jail population for each of Washington's 39 counties for calendar years 1995 to 2008. Thus, the balanced multiple time series panel dataset consists of 546 observations. From the Washington State Auditor, local jail expenditure data for counties were collected for 1993 to 2008, the earliest and latest years, as of winter 2010, available. The Auditor's data for the expenses includes all local jail expenditures (BARS code 527) except local probation costs (BARS code 527.40). These nominal annual dollar amounts were adjusted to 2009 dollars (JAILREAL) using the United States Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce, as published and forecasted by the Washington Economic and Revenue Forecast Council. The average daily jail population data (JAILADP) was obtained from the Washington Association of Sheriffs and Police Chiefs.

We computed the statewide average cost per jail ADP (in 2009 dollars) and plotted the results.

Exhibit D2.f
Average County Jail ADP Costs, 2009 Dollars
Fiscal Years 1993 to 2008



Over the entire 1993 to 2008 timeframe, the average statewide cost is \$28,900 per ADP, in 2009 dollars. Over these years, there has been an upward trend in the inflation-adjusted costs. We computed an estimate of the average annual real escalation rate in costs by estimating a linear line (shown on the chart) for this series. From this line, we computed the predicted values for 1993 (\$23,897) and 2008 (\$33,035) and calculated the average escalation rate for the 15 years, using equation (1), where FV is the 2008 estimated cost, PV is the 1993 estimate, and N is 15 years.

The annual rate of escalation is .022. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a.

To estimate the marginal annual operating costs of county jails, we conducted a time-series analysis of the panel data for Washington's 39 counties for 1993 to 2008. Thus the balanced panel includes a total of 546 observations. First, we tested each data series (JAILADP and JAILREAL) for unit roots. If unit roots are present, then a simple regression in levels of two unit root series can produce spurious results.⁴² We tested for unit roots with a panel unit root test, the Im, Pesaran, and Shin (IPS) test for individual unit root processes.

- For the JAILREAL expenditure series, the test without time trends failed to reject the null hypotheses that the series has a unit root (IPS p-value of 1.00). With time trends included, the IPS test continued to indicate a unit root (p-value .713). In first-differences, the test indicated a lack of a unit root (IPS p-value .000).
- For the JAILADP series, the IPS test without time trends failed to reject the null hypotheses that the series has a unit root (IPS p-value of .975). With time trends included, the IPS test continued to indicate a unit root (p-value .582). In first-differences, the test indicated a lack of a unit root (IPS p-value .000).
- With the IPS test indicating unit roots in both JAILREAL and JAILADP series, and no unit roots in first-differences, we proceeded to construct a model in first-differences.

⁴² Ibid., p. 636.

Since the two series have unit roots, we tested to determine if the two series together are cointegrated.⁴³ We used two versions of a panel cointegration test in EViews. Both the Pedroni Engle-Granger test (p-value .000) and the Kao Engle-Granger test (p-value .000) rejected the null hypothesis of no cointegration. We concluded that the two series together are I(0) cointegrated.

Since the two unit root series are cointegrated, we estimated an error correction model in first-differences. We tested alternative lag specifications of the JAILADP variable and concluded that three lags were appropriate. For the error correction term, we computed a cointegrating parameter from a simple model of: $JAILREAL = a + b(JAILADP)$.

The sum of the three ADP variables was \$21,469. The F-test of joint significance for the three ADP variables is marginally significant with a p-value of .113. The short-run marginal cost from the regression is the first lag term (\$3,457). We included cross-section (county) and period (year) fixed effects in the specification. We also included a lagged dependent variable on the right-hand side. Without this variable, the sum of the three ADP coefficients totaled \$37,637, an amount that seemed much higher than we expected. Thus, we included the lagged dependent variable in the model.⁴⁴

Exhibit D2.g

| Dependent Variable: JAILREAL-JAILREAL(-1) | | | | |
|--|-------------|-----------------------|-------------|--------|
| Method: Panel Least Squares | | | | |
| Date: 01/21/10 Time: 14:36 | | | | |
| Sample (adjusted): 1995 2008 | | | | |
| Periods included: 14 | | | | |
| Cross-sections included: 39 | | | | |
| Total panel (balanced) observations: 546 | | | | |
| White diagonal standard errors & covariance (d.f. corrected) | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | -682109.7 | 264036.1 | -2.583395 | 0.0101 |
| JAILREAL(-1)-JAILREAL(-2) | 0.359767 | 0.089133 | 4.036304 | 0.0001 |
| JAILADP-JAILADP(-1) | 3456.648 | 3050.223 | 1.133244 | 0.2577 |
| JAILADP(-1)-JAILADP(-2) | 8348.148 | 6128.536 | 1.362177 | 0.1738 |
| JAILADP(-2)-JAILADP(-3) | 9663.879 | 4591.016 | 2.104954 | 0.0358 |
| JAILREAL(-1)-39640.36*JAILADP(-1) | -0.266495 | 0.089148 | -2.989351 | 0.0029 |
| Effects Specification | | | | |
| Cross-section fixed (dummy variables) | | | | |
| Period fixed (dummy variables) | | | | |
| R-squared | 0.683040 | Mean dependent var | 439983.7 | |
| Adjusted R-squared | 0.646742 | S.D. dependent var | 2286829. | |
| S.E. of regression | 1359189. | Akaike info criterion | 31.18121 | |
| Sum squared resid | 9.03E+14 | Schwarz criterion | 31.63038 | |
| Log likelihood | -8455.470 | Hannan-Quinn criter. | 31.35680 | |
| F-statistic | 18.81750 | Durbin-Watson stat | 2.024971 | |
| Prob(F-statistic) | 0.000000 | | | |

Jail Capital Costs. Local Adult Jail capital costs for new beds were estimated from an informal internet review of current estimates for a variety of new jails around the country. We placed the estimate at \$150,000 capital cost per county jail bed. In our crime model, the total capital cost per bed is converted to an annualized capital payment, with equation (2), assuming a 25-year financing term (n), the bond financing rate entered in the model (i), and setting PV equal to the capital cost per bed converted to the base-year dollars chosen for the model.

⁴³ Ibid., p. 639.

⁴⁴ We also ran the preferred model shown above, but without the error correction. The coefficients from the three ADP variables totaled \$44,980—again, this sum seems too large based on prior expectations.

Local Juvenile Detention and Probation Per-Unit Costs

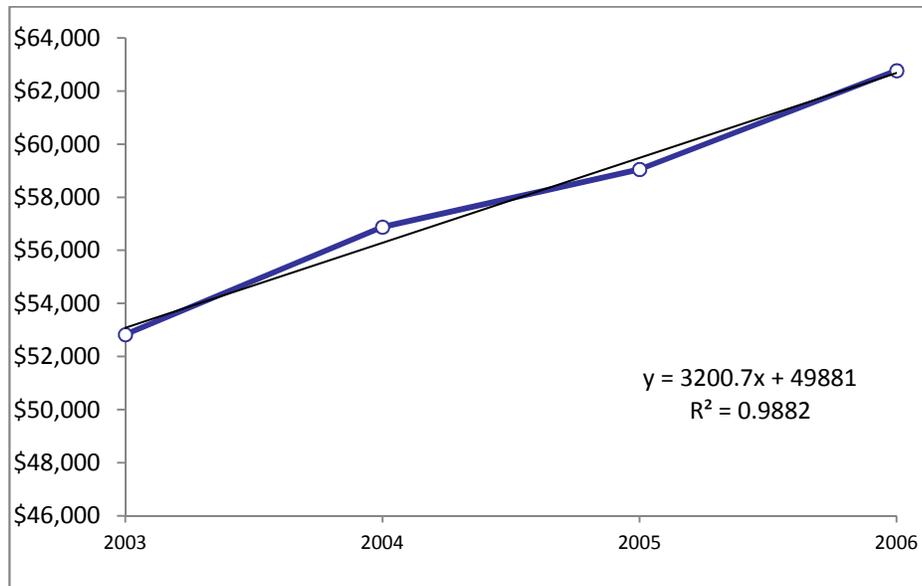
This section describes the steps we use to estimate marginal annual detention operating costs, and the long-run rate of real (inflation-adjusted) change in these costs of county-run juvenile detention facilities in Washington. We also describe our estimate of the capital cost per detention bed, as well as our estimate for the marginal annual costs of local juvenile probation and the real rate of annual escalation in these costs. All of these estimated cost parameters are entered into the crime model, as shown in Exhibit D2.a.

Detention Operating Costs. For an estimate of the marginal operating cost of state juvenile offender institutions, we conducted a time-series analysis of annual data for detention expenditures and average daily admissions to juvenile detention facilities in Washington. From the Washington State Auditor, local juvenile detention operating expenditure data for counties were collected for 1993 to 2008, the earliest and latest years available, as of winter 2010. The Auditor's data for the expenses include the categories for residential care&custdy-juvenilesvc (BARS 527.60) and juvenile facilities (BARS 527.80). Unfortunately, visual inspection of these historical data revealed significant problems and gaps, apparently caused by inconsistent reporting. We concluded that a consistent series could only be used for four years, 2003 to 2006. These nominal annual dollar amounts were adjusted to 2009 dollars using the United States Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce, as published and forecasted by the Washington Economic and Revenue Forecast Council.

To our knowledge, there is not a consistent statewide data series available for the average daily population of the county juvenile detention facilities. Instead, we collected annual admission data for the juvenile facilities; this information is collected and published by the Washington State Governor's Juvenile Justice Advisory Committee. From other data we have analyzed previously, it appears the average length of stay of a juvenile detention admission is about 12 days. Using this figure, along with the actual admission data, we estimated the average daily population (ADP) of the facilities statewide.

We computed the average costs per institutional ADP (in 2009 dollars) and plotted these data on this chart.

Exhibit D2.h
Average Local Juvenile Detention ADP Costs,
2009 Dollars, Fiscal Years 2003 to 2006



Over the 2003 to 2006 timeframe, the average annual cost is \$57,727 per ADP, in 2009 dollars. Over these years, there has been an upward trend in the inflation-adjusted costs. We computed an estimate of the average annual real escalation rate in costs by estimating a linear line (shown in the chart) for this series. From this line, we computed the predicted values for 2003 (\$53,131) and 2006 (\$62,742) and calculated the average escalation rate for the three years, using formula (1), where FV is the 2006 estimated cost, PV is the 2003 estimate, and N is three years.

The annual rate of real escalation is .057. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a. Because this is a high escalation rate, it will be important to seek additional information for this parameter.

Next, to estimate the marginal annual operating costs of police agencies, we conducted a time-series analysis of the panel data for Washington's 39 counties for 2003 to 2006. Because of the reasons mentioned above regarding the lack of a longer time series, we could not conduct unit root tests for these data. Since a regression in levels indicated a very high R-squared, and this often can indicate unit roots, and since so many of our other analyses of criminal justice data have revealed unit roots, we proceeded to construct a first-difference regression model.

We tested alternative lag specifications of the admission data. Our preferred model contained two lags and also a lagged dependent variable. Because of the lagging and, unfortunately, the already short time series, the model only had two periods for the 20 counties in Washington with juvenile detention facilities. The sum of the two admission coefficients is \$667. We converted this to an estimate of the annual marginal cost per ADP by, again, assuming a 12-day average length of stay. The result was an estimate of \$20,293 per annual ADP for juvenile detention marginal operating expenditures, in 2009 dollars. The following are the regression results obtained to support these calculations.

Exhibit D2.i

| Dependent Variable: JUVDETREAL-JUVDETREAL(-1) | | | | |
|---|-------------|-----------------------|-------------|----------|
| Method: Panel Least Squares | | | | |
| Date: 02/05/10 Time: 17:16 | | | | |
| Sample (adjusted): 2005 2006 | | | | |
| Periods included: 2 | | | | |
| Cross-sections included: 20 | | | | |
| Total panel (balanced) observations: 40 | | | | |
| White cross-section standard errors & covariance (d.f. corrected) | | | | |
| WARNING: estimated coefficient covariance matrix is of reduced rank | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 80820.93 | 8253.006 | 9.792908 | 0.0000 |
| JUVDETREAL(-1)-JUVDETREAL(-2) | -0.139491 | 0.082108 | -1.698865 | 0.0980 |
| JUVDETADM-JUVDETADM(-1) | 445.0912 | 246.1837 | 1.807964 | 0.0790 |
| JUVDETADM(-1)-JUVDETADM(-2) | 222.0772 | 57.98376 | 3.829989 | 0.0005 |
| R-squared | 0.087247 | Mean dependent var | | 44115.96 |
| Adjusted R-squared | 0.011185 | S.D. dependent var | | 333851.7 |
| S.E. of regression | 331979.4 | Akaike info criterion | | 28.35817 |
| Sum squared resid | 3.97E+12 | Schwarz criterion | | 28.52706 |
| Log likelihood | -563.1635 | Hannan-Quinn criter. | | 28.41924 |
| F-statistic | 1.147044 | Durbin-Watson stat | | 2.026817 |
| Prob(F-statistic) | 0.343320 | | | |

Local Detention Capital Costs. Per-bed capital costs for a new detention facility would run \$200,000 per bed.⁴⁵ In our crime model, the total capital cost per arrest is converted to an annualized capital payment, with equation (2), assuming a 25-year financing term (n), the bond financing rate entered in the model (i), and setting PV equal to the capital cost per bed converted to the base-year dollars chosen for the model.

Local Juvenile Probation Per-Unit Costs

We searched for longitudinal time-series data to estimate the average annual cost of county-run juvenile probation services in Washington. Unfortunately, we did not locate a consistent set of expenditure information or average daily caseload information that would have allowed us to perform a valid time-series analysis. The expenditure data from the Washington State Auditor contain a considerable number of county jurisdictions that do not report, every year, their juvenile court expenditures. And, as far as we know, there is not a data source for the average daily juvenile court probation caseloads in Washington.

⁴⁵ Capital costs for a typical new local juvenile detention facility were estimated from personal communication with Washington's Juvenile Rehabilitation Administration staff.

Therefore, we estimated marginal juvenile court probation costs with the following procedures.

- From the State Auditor, we collected statewide juvenile court probation expenditure data for calendar year 2008, the latest year reported as of March 2010. These data appear to be reasonably complete with the exception of Snohomish County that did not report juvenile county probation expenditures that year. The total reported expenditures for juvenile probation for the state was \$29,203,723 for 2008. Again, this figure does not include Snohomish County.
- From the Administrative Office of the Courts, we collected the reported number of juvenile court community supervision sentences and sentences with detention and community supervision for 2008. The total was 5,660.
- From an Institute survey of juvenile court activities in 1995, we calculated that the average length of stay on juvenile court probation in Washington is 6.8 months.⁴⁶
- We then estimated the 2008 average daily probation caseload of juvenile courts as 3,207 (5,660 times 6.8 divided by 12 months).
- We adjusted the statewide average daily caseload to remove Snohomish County by subtracting an estimate of Snohomish's average daily caseload. Snohomish had 705 juvenile court community supervision sentences and sentences with detention and community supervision in 2008. An estimate of the average daily caseload in Snohomish for 2008 was 400 (705 times 6.8 divided by 12 months), assuming the same 6.8-month average length of stay on juvenile court probation. Thus, after removing Snohomish, an estimate of the adjusted statewide average daily probation caseload was 2,808 in 2008.
- We then computed the average expenditure per average annual daily caseload to be \$10,401 (\$29,203,723 divided by 2,808).
- From this estimate of the *average* expenditure per average annual caseload, we estimated the *marginal* expenditure per average annual caseload. We found from our time-series analysis of the community supervision costs of the Department of Corrections that the ratio of marginal costs to average costs was .50 (see local community supervision section where marginal DOC community supervision costs are estimates as \$1,861 and average costs are \$3,707). Multiplying \$10,401 by .50 provides an estimate, \$5,200 in 2008 dollars, of the marginal cost per average annual juvenile court caseload. This estimate is included as a parameter in the crime model, as shown in Exhibit D2.a.

State Juvenile Rehabilitation Administration (JRA) Per-Unit Costs

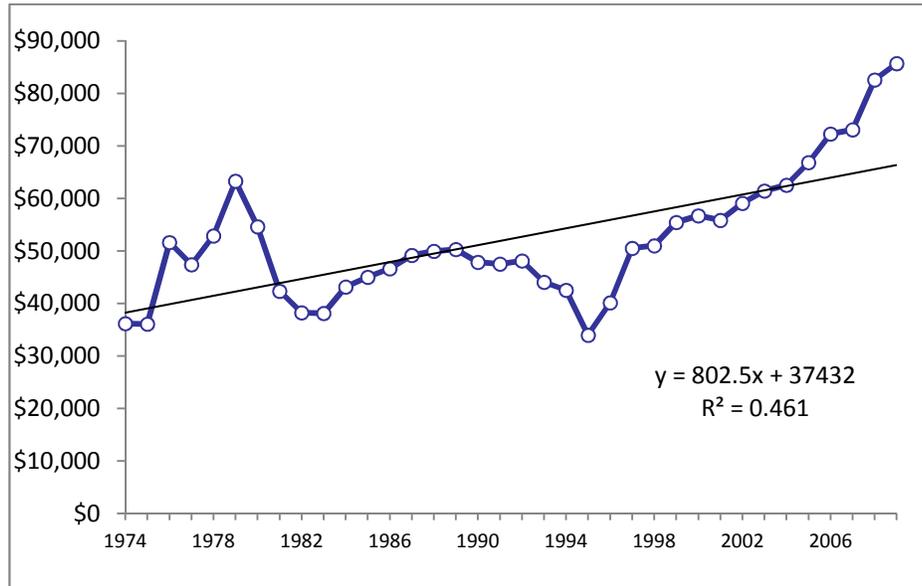
This section describes the steps we use to estimate marginal annual institution operating costs, and the long-run rate of real (inflation-adjusted) change in these costs, of the Washington State Juvenile Rehabilitation Administration (JRA). JRA is Washington's state juvenile justice agency; juvenile offenders are sentenced to JRA based on Washington's sentencing laws and practices. We also describe our estimate of the JRA capital cost per institutional bed as well as our estimate for the marginal annual costs of community supervision for juvenile parole supervision in Washington, and the real rate of annual escalation in these costs. All of these estimated cost parameters are entered into the crime model, as shown in Exhibit D2.a.

Institutional Operating Costs. For an estimate of the marginal operating costs of state juvenile offender institutions, we conducted a time-series analysis of annual data for institutional expenditures and average daily institutional population for JRA for fiscal years 1974 to 2009. The expenditure data were obtained from the Washington State's Legislative Evaluation and Accountability Program (LEAP) for Agency 300 (Juvenile Rehabilitation Administration) for code 2000 (institutional services). The LEAP data series for JRA begins in fiscal year 1974. We converted the expenditure data to 2009 dollars (JRAREAL) using the United States Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce, as published and forecasted by the Washington Economic and Revenue Forecast Council. The average daily population for JRA institutions (JRAADP) series is from the Washington State Caseload Forecast Council for Fiscal Years 1997 to 2009, with data from 1974 to 1996 taken from annual reports of the Governor's Juvenile Justice Advisory Committee and data from various issues of the Databook series published by the Washington State Office of Financial Management.

We computed the average costs per institutional ADP (in 2009 dollars) and plotted these data in Exhibit D2.j.

⁴⁶ Burley, M., & Barnoski, R. (1997, April). *Washington State juvenile courts: Workloads and costs* (Document No. 97-04-1201). Olympia: Washington State Institute for Public Policy, Table 2.

Exhibit D2.j
Average JRA Institution ADP Costs, 2009 Dollars
Fiscal Years 1974 to 2009



Over the entire 1974 to 2009 timeframe, the average cost is \$51,716 per ADP, in 2009 dollars. Over these years, there has been an upward trend in the inflation-adjusted costs. We computed an estimate of the average annual real escalation rate in costs by estimating a linear line (shown in the chart) for this series. From this line, we computed the predicted values for 1974 (\$38,274) and 2009 (\$66,379) and calculated the average escalation rate for the 35 years, using formula (1), where FV is the 2009 estimated cost, PV is the 1974 estimate, and N is 35 years.

The annual rate of escalation is .016. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a. The data plotted on the chart reveals that in the last five years, the growth in real average costs has been on a steeper incline compared with the annual growth rate over the entire period of record. Thus, our estimate of .016 may be on the low side of recent trends persist.

To estimate the marginal annual operating cost of a state institutional bed, we conducted a time-series analysis of these data. First, we tested each data series (JRAADP and JRAREAL) for unit roots. If unit roots are present, then a simple regression in levels of two unit root series can produce spurious results.⁴⁷ We tested for unit roots with an Augmented Dickey-Fuller (ADF) test.

- For the JRAREAL expenditure series, the ADF test failed to reject the null hypotheses that the series has a unit root, with p-values of .511 without a time trend and .620 with a time trend, indicating a unit root with both tests. In first-differences, on the other hand, the ADF p-value for the JRAREAL series is 0.000.
- For the JRAADP series, the p-values were .299 without a time trend and .760 with a time trend, indicating a unit root in both tests. In first-differences, the ADF p-value for the JRAADP series is 0.049.
- With both JRAREAL and JRAADP series indicating unit roots in levels and no unit roots in first-differences, we proceeded to construct a model in first-differences.

Since the two series have unit roots, we tested to determine if the two series together are cointegrated.⁴⁸ We used an Engle-Granger test to determine whether the residuals from a cointegrating regression of the two series are integrated of order 1 (i.e., I(1), unit root). The resulting tau-statistic from the regression was -1.03, which is well below the Engle-Granger critical value of -3.9 (p-value .01) for the null hypothesis that the residual series has a unit root. Therefore, this test did not reject the null hypothesis that the two series, jointly, have a unit root. We concluded that the two series together are I(1) and, therefore, not I(0) cointegrated.

⁴⁷ Wooldridge, 2009, p. 636.

⁴⁸ Ibid., p. 639.

We then computed a first-difference model with three lags on the first-differenced JRAADP variables and obtained the following result:

Exhibit D2.k

| Dependent Variable: JRAREAL-JRAREAL(-1) | | | | |
|---|-------------|-----------------------|-------------|----------|
| Method: Least Squares | | | | |
| Date: 01/20/10 Time: 15:53 | | | | |
| Sample (adjusted): 1975 2009 | | | | |
| Included observations: 35 after adjustments | | | | |
| HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000) | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | -321480.3 | 928044.0 | -0.346406 | 0.7315 |
| JRAADP-JRAADP(-1) | 5845.823 | 16565.04 | 0.352901 | 0.7266 |
| JRAADP(-1)-JRAADP(-2) | 28438.73 | 18767.99 | 1.515279 | 0.1402 |
| JRAADP(-2)-JRAADP(-3) | 2458.799 | 13179.94 | 0.186556 | 0.8533 |
| RPCI(-1)-RPCI(-2) | 2276.323 | 888.6560 | 2.561534 | 0.0157 |
| R-squared | 0.257160 | Mean dependent var | | 1038534. |
| Adjusted R-squared | 0.158115 | S.D. dependent var | | 5199909. |
| S.E. of regression | 4771140. | Akaike info criterion | | 33.72563 |
| Sum squared resid | 6.83E+14 | Schwarz criterion | | 33.94783 |
| Log likelihood | -585.1986 | Hannan-Quinn criter. | | 33.80233 |
| F-statistic | 2.596387 | Durbin-Watson stat | | 2.090018 |
| Prob(F-statistic) | 0.056213 | | | |

After testing different model specifications, our preferred model includes three lagged first-difference JRAADP variables and a first-differenced covariate (RPCI, real per capita income). We examined multiple lags in the JRAADP variables and three lags seemed appropriate. The sum of the three lagged coefficients was \$36,743, in 2009 dollars. This is our estimate of the marginal operating cost of an annual JRA bed.⁴⁹ The three ADP variables were jointly significant with a p-value on the F test of .0473. The short-run marginal cost from the regression is the first lag term (\$5,846).

JRA Capital Costs. JRA capital costs for typical new institutional beds were estimated from personal communication with JRA staff. Per-bed capital costs for a new medium secure facility would run \$125,000 to \$175,000 per bed. In our crime model, the total capital cost per arrest is converted to an annualized capital payment, with equation (2), assuming a 25-year financing term (n), the bond financing rate entered in the model (i), and setting PV equal to the capital cost per bed converted to the base year dollars chosen for the model.

JRA Parole Costs. We were unable to obtain a long-term data set to analyze the marginal cost of JRA parole services. The electronic data for parole expenditures were only available starting in fiscal year 2000 and, beginning in fiscal year 2006, there was a significant accounting change that rendered the post-2005 data unusable for measuring parole expenditures. We do have consistent parole average daily population data from 1981 through 2009. We intend to obtain earlier expenditure data which may allow a regression analysis. In the meantime, we calculated an average parole cost by summing inflation-adjusted JRA parole costs from 2000 to 2005: \$43,004,688 (in 2009 dollars). The sum of the average daily parole caseloads during these same years was 5,481. Thus, the average annual expenditure per parole average daily population is \$7,847, in 2009 dollars. From this estimate of the average expenditure per average annual caseload, we estimated the marginal expenditure per average annual caseload. We found from our time-series analysis of the community supervision costs of the Department of Corrections that the ratio of marginal costs to average costs was .50. Multiplying \$7,847 by .50 provides an estimate, \$3,923 in 2009 dollars, of the marginal cost per average annual JRA parole caseload. This estimate is included as a parameter in the crime model, as shown in Exhibit D2.a.

⁴⁹ We also estimated a model identical to our preferred model but with a lagged first-differenced dependent variable on the right-hand side. The sum of the three ADP coefficients was \$39,138, only slightly larger than our preferred model. For purposes of our benefit-cost model, which is to compute benefit-cost estimates of evidence-based programs that reduce crime, our preference is to use the slightly more cautious estimate.

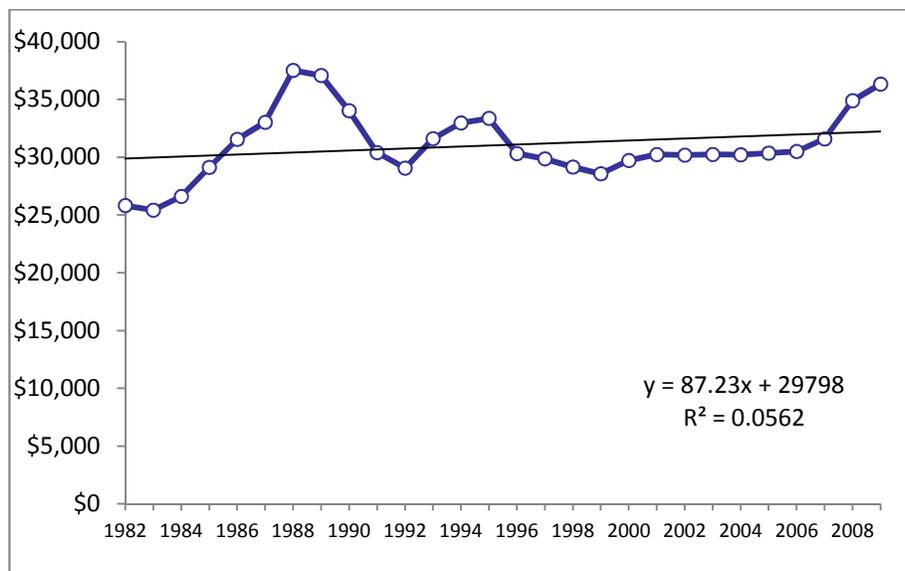
State Department of Corrections (DOC) Per-Unit Costs

This section describes the steps we used to compute estimates of Washington Department of Corrections' marginal annual prison operating costs and the long-run rate of change in these costs. We also provide our estimate of the capital cost of a prison bed. Additionally, we describe our estimate for the annual cost of community supervision for adult felony offenders in Washington, and the real rate of annual escalation in this cost.

Prison Operating Costs. For prison operating costs, we analyzed annual data for DOC institutional expenditures and average daily prison population for fiscal years 1982 to 2009. The expenditure data were obtained from Washington State's Legislative Evaluation and Accountability Program (LEAP) for Agency 310 (Department of Corrections) for code 200 (correctional expenditures); the LEAP data series for DOC begins in fiscal year 1982. The "correctional expenditures" category pertains to operating expenses for running the state's prison system, not the community corrections system. We converted the expenditure data to 2009 dollars using the United States Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce, as published and forecasted by the Washington Economic and Revenue Forecast Council. The average daily prison population (ADP) series is from the Washington State Caseload Forecast Council for fiscal years 1993 to 2009, with data for earlier years taken from various issues of the Databook series published by the Washington State Office of Financial Management.

We computed the average cost per prison ADP (in 2009 dollars) for 1982 to 2009 and plotted the results.

Exhibit D2.I
Average DOC ADP Prison Costs, 2009 Dollars
Fiscal Years 1982 to 2009



Over the entire 1982 to 2009 timeframe, the average cost is \$31,446 per ADP, in 2009 dollars. Over these years, there has been a slight upward trend in the inflation-adjusted per-unit costs, as revealed by the linear regression line shown in Exhibit D2.I. To compute an estimate of the long-run growth rate in real cost per average daily population, we calculated the predicted values from the regression line for 1982 (\$29,915) and 2009 (\$32,266) and calculated the annual rate of escalation for the 27 years using equation (1), where FV is the 2009 cost estimate, PV is the 1982 estimate, and N is 27 years.

The annual rate of real escalation in average costs is .003. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a.

To estimate marginal prison operating costs, we conducted a time-series analysis of total annual real operating costs (DOCREAL) and the total annual prison average daily population (DOCADP). First, we tested each data series for unit roots. If unit roots are present, then a simple regression in levels of two unit root series can produce spurious results.⁵⁰ We tested for unit roots with an Augmented Dickey-Fuller (ADF) test.

- For the DOCREAL expenditure series, the ADF test failed to reject the null hypotheses that the series has a unit root with p-values of .9999 without a time trend and .9978 with a time trend. In first-differences, on the other hand, the ADF p-value for the DOCREAL series was 0.0146, indicating a lack of a unit root in a first-differenced data series.
- For the DOCADP series, the p-values for the ADF test were .8668 without a time trend and .2744 with a time trend; both tests indicate that the DOCADP series in levels has a unit root. In first-differences, the ADF p-value for the DOCADP series was 0.0458 indicating a lack of a unit root in first-differences.
- With both DOCREAL and DOCADP series indicating unit roots in levels and no unit roots in first-differences, we proceeded to construct models in first-differences.⁵¹

Assuming the two series have unit roots, we tested to determine if the two series together are cointegrated.⁵² We used an Engle-Granger test to determine whether the residuals from the cointegrating regression were integrated of an order of 1 (i.e., I(1), a unit root). The resulting tau-statistic from the regression was -2.667, which is below the Engle-Granger critical value of -3.9 (p-value .01). Therefore, this test did not reject the null hypothesis that the two series, jointly, have a unit root. We concluded that the two series together are I(1) and, therefore, not cointegrated.

Exhibit D2.m

| Dependent Variable: DOCREAL-DOCREAL(-1) | | | | |
|---|-------------|-----------------------|-------------|----------|
| Method: Least Squares | | | | |
| Date: 01/18/10 Time: 16:09 | | | | |
| Sample (adjusted): 1983 2009 | | | | |
| Included observations: 27 after adjustments | | | | |
| HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 3.0000) | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 12790705 | 6212729. | 2.058790 | 0.0521 |
| DOCADP-DOCADP(-1) | 4495.187 | 6295.155 | 0.714071 | 0.4830 |
| DOCADP(-1)-DOCADP(-2) | -4288.905 | 5011.822 | -0.855758 | 0.4018 |
| DOCADP(-2)-DOCADP(-3) | 6745.884 | 3736.465 | 1.805419 | 0.0854 |
| DOCADP(-3)-DOCADP(-4) | 6968.766 | 2879.800 | 2.419879 | 0.0247 |
| RPCI(-1)-RPCI(-2) | 2355.135 | 3505.699 | 0.671802 | 0.5090 |
| R-squared | 0.128695 | Mean dependent var | | 21124103 |
| Adjusted R-squared | -0.078759 | S.D. dependent var | | 14953657 |
| S.E. of regression | 15531362 | Akaike info criterion | | 36.14775 |
| Sum squared resid | 5.07E+15 | Schwarz criterion | | 36.43571 |
| Log likelihood | -481.9946 | Hannan-Quinn criter. | | 36.23338 |
| F-statistic | 0.620356 | Durbin-Watson stat | | 1.290263 |
| Prob(F-statistic) | 0.685814 | | | |

Since the two unit root series are not jointly cointegrated, we did not estimate an error correction model and, instead, estimated a first-difference model.⁵³ The following results were obtained.

After testing different model specifications, our preferred model includes regressors with four lagged first-difference DOCADP variables and a first-differenced covariate (RPCI, real per capita income). We examined different numbers of lags in the DOCADP variables, and four lags seemed appropriate empirically and logically given our knowledge of state budgeting processes. The four DOCADP lags are jointly statistically significant (F test p-value .0085). The short-run marginal cost from the regression is the first lag term (\$4,495).

⁵⁰ Wooldridge, 2009, p. 636.

⁵¹ Ibid., p. 643.

⁵² Ibid., p. 639.

⁵³ Ibid., p. 643.

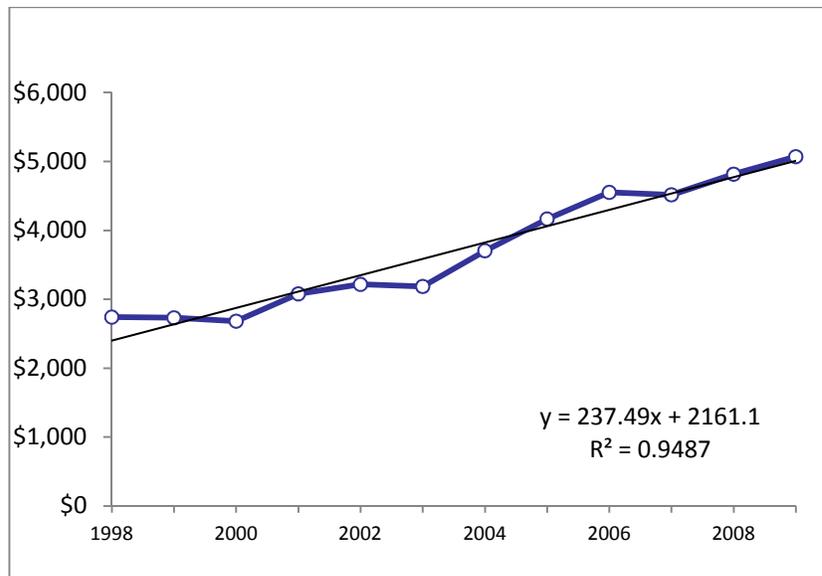
The sum of the four DOCADP distributed lags (the long-run multiplier) is \$13,921. This figure, \$13,921 per ADP (in 2009 dollars), represents our preferred estimate of the long-run incremental expenditures to DOC for a year in prison. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a.⁵⁴

The readily available annual time series for this analysis, unfortunately, was limited from 1982 to 2009, because expenditure data (DOCREAL) were only available from 1982 onward. We reviewed this marginal cost per prison ADP with legislative and executive fiscal staff to determine the accuracy of our estimate in the budgeting world. It was agreed upon that the marginal cost per prison ADP is \$12,722.

Prison Capital Costs. DOC capital costs for new institutional beds were estimated. Capital cost estimates for the relatively new Coyote Ridge medium security facility in Washington were obtained from legislative fiscal staff. The 2,048 bed facility cost \$232,118,000 (thus, a per-bed cost of \$113,339) and was completed in 2008. We recorded this per-bed cost figure as 2007 dollars since it is likely that was when most of the construction dollars were spent. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a. In our crime model, the total construction costs per-bed are converted to an annualized capital payment, with equation (2), assuming a 25-year financing term, the bond financing rate entered in the model, and setting PV equal to the per-bed construction cost converted to the base year dollars chosen for the model.

Community Supervision Operating Costs. We analyzed Department of Corrections' community supervision cost for all felony offenders on active supervision regardless of sentence type (prison or jail). For community supervision costs, we analyzed annual data for DOC community supervision expenditures and average daily community population for Fiscal Years 1998 to 2009. The expenditure data were obtained from Washington State's Legislative Evaluation and Accountability Program (LEAP) for Agency 310 (Department of Corrections) for code 300 (community supervision); the LEAP data series for DOC begins in fiscal year 1982. Community supervision population data were obtained from the Washington Caseload Forecast Council, which maintains data back to fiscal year 1998. We calculated annual cost per average daily community population and converted to 2009 dollars using the aforementioned price index. The average community supervision cost over the 1998 to 2009 period is \$3,657.

Exhibit D2.n
Average DOC Average Daily Community Supervision Costs,
2009 Dollars, Fiscal Years 1998 to 2009



⁵⁴ As an additional test, we ran our preferred model with a lagged first difference dependent variable on the right-hand side of the equation. The results were somewhat similar to our preferred model (e.g., the sum of the three positive lagged DOCADP coefficient was \$15,413, but the three coefficient together were only marginally significant with a F-test p-value of .1111).

Over the 1998 to 2009 period, there was a significant upward trend in the inflation-adjusted per-unit costs, as revealed by the linear regression line shown in Exhibit D2.n. To compute an estimate of the long-run growth rate in real cost per average daily population, we calculated the predicted values from the regression line for 1998 (\$2,399) and 2009 (\$4,773) and calculated the annual rate of escalation for the 11 years using equation (1) where FV is the cost estimate for 2009, PV is the estimate for 1998, and N is 11 years.

The annual rate of real escalation in average costs is 0.064. This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a. This estimate seems high, and it will be useful to monitor actual expenditure trends in the years ahead.

To estimate marginal community supervision operating costs, we conducted a time-series analysis of total annual real operating costs (DOCCSREAL) and the total annual community supervision average daily population (DOCCSADP). First, we tested each data series for unit roots. If unit roots are present, then a simple regression in levels of two unit root series can produce spurious results.⁵⁵ We tested for unit roots with an Augmented Dickey-Fuller (ADF) test.

- For the DOCCSREAL expenditure series, the ADF test failed to reject the null hypotheses that the series has a unit root with p-values of .8446 without a time trend, but was significant at .0276 with a time trend. In first-differences, on the other hand, the ADF p-value for the DOCCSREAL series was 0.0263, indicating a lack of a unit root in a first-differenced data series.
- For the DOCCSADP series, the p-values for the ADF test were .2243 without a time trend and .2682 with a time trend; both tests indicate that the DOCCSADP series in levels has a unit root. In first-differences, the ADF p-value for the DOCCSADP series was 0.1318 indicating, marginally, a lack of a unit root in first-differences.
- With both DOCCSREAL and DOCCSADP series indicating, generally, unit roots in levels (with the exception of an ADF test with a time trend for DOCCSREAL) and, marginally, no unit roots in first-differences, we proceeded to construct models in first-differences. We also tested models in levels.

Assuming the two series have unit roots, we tested to determine if the two series together are cointegrated.⁵⁶ We used an Engle-Granger test to determine whether the residuals from a cointegrating regression of the two series are integrated of order 1 (i.e., I(1), a unit root). The resulting tau-statistic from the regression was -1.45, which is well below the Engle-Granger critical value of -3.9 (p-value .01) for the null hypothesis that the residual series has a unit root. Therefore, this test did not reject the null hypothesis that the two series, jointly, have a unit root. We concluded that the two series together are I(1) and, therefore, not cointegrated.

Since the two unit root series are not jointly cointegrated, we did not estimate an error correction model. Since there was some ambiguity over the existence of unit roots, we ran a basic regression in both levels and first-differences. The following first-difference results, our preferred approach, were obtained. The sum of the three coefficients total \$1,861 per ADP, in 2009 dollars.⁵⁷ This point estimate is included as a parameter in the crime model, as shown in Exhibit D2.a. The three ADP variables are jointly significant with a p-value on the f-test of .0042.

⁵⁵ Wooldridge, 2009, p. 636.

⁵⁶ Ibid., p. 639.

⁵⁷ We ran this same model with a lagged first difference dependent variable on the right-hand side and the sum of the three coefficients was \$2,407. For purposes of our benefit-cost model, which is to compute benefit-cost estimates of evidence-based programs that reduce crime, our preference is to use the non-lagged dependent variable model since it produces a slightly more cautious estimate.

Exhibit D2.o

| Dependent Variable: DOCCSREAL-DOCCSREAL(-1) | | | | |
|---|-------------|-----------------------|-------------|----------|
| Method: Least Squares | | | | |
| Date: 01/19/10 Time: 16:50 | | | | |
| Sample (adjusted): 2001 2009 | | | | |
| Included observations: 9 after adjustments | | | | |
| HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 3.0000) | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 9209858. | 1150172. | 8.007374 | 0.0005 |
| DOCCSADP-DOCCSADP(-1) | 1193.120 | 220.3772 | 5.413988 | 0.0029 |
| DOCCSADP(-1)-DOCCSADP(-2) | 449.9942 | 659.9840 | 0.681826 | 0.5256 |
| DOCCSADP(-2)-DOCCSADP(-3) | 217.7877 | 483.4093 | 0.450525 | 0.6712 |
| R-squared | 0.542175 | Mean dependent var | | 8708889. |
| Adjusted R-squared | 0.267480 | S.D. dependent var | | 5067302. |
| S.E. of regression | 4336970. | Akaike info criterion | | 33.70435 |
| Sum squared resid | 9.40E+13 | Schwarz criterion | | 33.79201 |
| Log likelihood | -147.6696 | Hannan-Quinn criter. | | 33.51519 |
| F-statistic | 1.973736 | Durbin-Watson stat | | 2.347624 |
| Prob(F-statistic) | 0.236419 | | | |

This first-difference model is our preferred model. Our model in levels revealed a negative relationship between community supervision average daily population and real expenditures, which does not make intuitive budgeting sense. The first-difference model, shown above, produced the most plausible estimates, given our knowledge of state budget processes.

Victimizations Per-Unit Cost

In addition to costs paid by taxpayers, many of the costs of crime are borne by victims. Some victims lose their lives. Others suffer direct, out-of-pocket, personal or property losses. Psychological consequences also occur to crime victims, including feeling less secure in society. The magnitude of victim costs is very difficult—and in some cases impossible—to quantify.

In recent years, however, analysts have taken significant steps in estimating crime victim costs. We use a consistent set of estimates (McCollister, 2010), with some modifications, in the Institute's benefit-cost model.⁵⁸ These crime victim costs build on and modify the previous work prepared for the U.S. Department of Justice by Miller, Cohen, and Wiersema (1996).⁵⁹

These studies divide crime victim costs into two types:

- a) *Tangible* victim costs, which include medical and mental health care expenses, property damage and losses, and the reduction in future earnings incurred by crime victims; and
- b) *Intangible* victim costs, which place a dollar value on the pain and suffering of crime victims. In these two studies, the intangible victim costs are computed, in part, from jury awards for pain, suffering, and lost quality of life.

The McCollister study divides total tangible costs of crime into tangible victim costs, criminal justice system costs, and crime career costs of offenders (estimates of the economic productivity losses for offenders). In the Institute's model, we only include McCollister's tangible victim costs since we estimate criminal justice costs separately. We currently do not make estimates of the crime career costs of offenders.

We also use McCollister's intangible victim costs with one exception. McCollister computes a "corrected risk-of-homicide cost" as part of crime specific intangible victim costs. This is done because, according to McCollister, the FBI's Uniform Crime Reports (UCR) classifies some homicides as other non-homicide crimes when certain offense information is

⁵⁸ McCollister, K. E., French, M. T., & Fang, H. (2010). The cost of crime to society: New crime-specific estimates for policy and program evaluation. *Drug and Alcohol Dependence*, 108(1), 98-109.

⁵⁹ Miller, T. R., Cohen, M. A., & Wiersema, B. (1996). *Victim costs and consequences: A new look* (Document No. NCJ 155282). Washington, DC: National Institute of Justice.

lacking. This FBI reporting practice requires the adjustment made by McCollister. For application to the Institute's benefit-cost model, however, this adjustment is not necessary. The Institute's crime cost estimates are applied to accurately classified conviction data from Washington State; convictions for homicide are not misclassified as other crimes in the Washington system. See section D2.3 of this Appendix for a description of the Institute's data sources for counting convictions.

The Institute's model also has one crime category for felony property crimes. The McCollister study breaks the Institute's property crime classification into motor vehicle theft, household burglary, and larceny/theft. We use these three categories and compute a weighted average property category using the estimated number of crimes calculated for Washington as weights.

The Institute's modified McCollister crime victim cost estimates are included in the crime model, as shown in Exhibit D2.a.

Not all crime is reported to, or acted upon by, the criminal justice system. When crimes are reported by citizens or detected by police or other officials, however, the use of taxpayer-financed resources begins. The degree to which these resources are used depends on the crime as well as the policies and practices governing the criminal justice system's response. In the preceding section, we describe the *per-unit* marginal cost estimates used in our model. In this section, we discuss *how many units* of the criminal justice system are used when a crime occurs.

Once a person is convicted for a criminal offense, sentencing policies and practices in Washington affect the use of different local and state criminal justice resources. Exhibit D2.p is a screen shot from the Institute's benefit-cost model that displays how criminal justice resources in Washington State are used in response to crime. The estimates for each row in Exhibit D2.p are described below.

Probability of Resource Use. The first block of information in Exhibit D2.p displays parameters indicating the probability that a person convicted for one of the seven crime categories modeled will receive a sentence to a juvenile state institution (instead of local juvenile detention) or adult state prison (instead of local adult jail). For example, if an adult offender is convicted of robbery, there is a 71 percent chance the offender will receive a prison sentence and a 29 percent chance of receiving a jail sentence. These sentencing probabilities were obtained from the Washington State Sentencing Guidelines Commission.⁶⁰

Number of Years of Use Per Resource. We estimate the average number of years various criminal justice resources are used for each of the crime categories.

Juvenile Detention (with local or state sentence). Unfortunately, Washington does not have an annual reporting system on local juvenile detention length of stay. Therefore, the average length of stay at local juvenile detention facilities and the average length of local probation were estimated from an earlier survey of juvenile courts conducted by the Institute.⁶¹

Juvenile Local Supervision. The average length of stay on probation was also estimated from the same survey of juvenile courts conducted by the Institute.⁶²

Juvenile State Institution. The average length of stay in a juvenile state institution was estimated using data obtained from the Sentencing Guidelines Commission.⁶³

Juvenile State Supervision. The average length of stay on juvenile parole was estimated using information obtained from the Juvenile Rehabilitation Administration.⁶⁴

Adult Jail, With Local Sentence. The average length of stay in jail for local sentences was estimated using data from the Sentencing Guidelines Commission.⁶⁵

Adult Jail, With Prison Sentence. Analysis from the Department of Corrections on the credit for time served in jail was used to estimate the total length of stay in jail prior to prison.⁶⁶

⁶⁰ Juvenile sentencing information obtained from SGC staff via email on March 10, 2010. Adult sentencing information obtained from: Sentencing Guidelines Commission. (2009, January). *Statistical summary of adult felony sentencing: Fiscal year 2008*. Olympia, WA: Author, Table 1.

⁶¹ Burley & Barnoski, 1997.

⁶² Ibid.

⁶³ Washington State Sentencing Guidelines Commission (personal communication, March 10, 2010).

⁶⁴ Washington State Juvenile Rehabilitation Administration (personal communication, April 18, 1997).

⁶⁵ Sentencing Guidelines Commission, 2009, Table 1.

⁶⁶ Washington State Department of Corrections (personal communication, November 7, 1996).

Adult Community Supervision and Adult Post Prison Supervision. These numbers were obtained from the Sentencing Guidelines Commission.⁶⁷

Adult Prison. The information for the average sentence received for adults sentenced to a state prison comes from Sentencing Guidelines Commission data. As a result of good-time reductions to some prison sentences, the average time actually served is often shorter than the original sentence. Exhibit D2.p shows the average prison length of stay, which is computed in the model by multiplying the sentence length of stay by an average percentage good-time reduction. The data on the average sentence reductions, by crime, were obtained from an analysis supplied by the Washington State Department of Corrections.

Exhibit D2.p

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

General | Economic | **Crime** | Education | Child Welfare | Substance Use | Health Care | Mental Health | Public Asst | Housing | Teen Birth | Outcomes & Links

Crime

Back to Main Model

Per Unit Costs | Resource Use | Offender Populations | Victimization | Program Participation

Total current prison average daily population (ADP): 18400

| | Felonies Violent & Property Crime Categories | | | | | Other | | Sums | | Year of Data |
|---|--|-------------------|---------|--------------------|-----------------|-------------|-------------|---------------|--------------------------|--------------|
| | Murder | Felony Sex Crimes | Robbery | Aggravated Assault | Felony Property | Felony Drug | Misdemeanor | Total Violent | Total Violent & Property | |
| Probability of Resource Use | | | | | | | | | | |
| Juvenile State Institution | 0.86 | 0.46 | 0.68 | 0.34 | 0.15 | 0.14 | 0.02 | 0.43 | 0.24 | 2009 |
| Adult State Prison | 0.96 | 0.71 | 0.72 | 0.39 | 0.35 | 0.30 | 0.00 | 0.49 | 0.41 | 2009 |
| Juvenile Local Supervision | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 | 1.00 | |
| Juvenile State Supervision (Parole) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 | 1.00 | |
| Adult Community Supervision Post-Jail | 1.00 | 0.85 | 0.89 | 0.69 | 0.17 | 0.73 | 0.00 | 0.73 | 0.38 | 2008 |
| Adult Community Supervision Post-Prison | 0.99 | 0.96 | 0.98 | 0.81 | 0.29 | 0.96 | 0.00 | 0.87 | 0.61 | 2008 |
| Number of Years of Use Per Resource | | | | | | | | | | |
| Juvenile Detention, with Local Sentence | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.00 | | | 2008 |
| Juvenile Detention, with State Sentence | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.00 | | | 1996 |
| Juvenile Local Supervision | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.00 | | | 1996 |
| Juvenile State Institution | 1.65 | 0.90 | 0.96 | 0.67 | 0.53 | 0.63 | 0.19 | | | 2009 |
| Juvenile State Supervision | 0.47 | 1.49 | 0.44 | 0.45 | 0.48 | 0.55 | 0.47 | | | 2009 |
| Adult Jail, with Local Sentence | 0.74 | 0.59 | 0.55 | 0.36 | 0.23 | 0.23 | 0.10 | 0.39 | 0.29 | 2009 |
| Adult Jail, with Prison Sentence | 1.08 | 0.48 | 0.44 | 0.37 | 0.32 | 0.28 | 0.00 | | | 2009 |
| Adult Community Supervision, Jail Sentence | 2.00 | 2.50 | 1.01 | 0.82 | 0.24 | 0.90 | 0.50 | 1.17 | 0.92 | 2008 |
| Adult Prison | 14.84 | 6.06 | 3.95 | 2.64 | 1.65 | 1.35 | 0.00 | 4.35 | 2.99 | 2009 |
| Adult Community Supervision, Post-Prison | 3.91 | 3.70 | 2.94 | 1.67 | 0.51 | 1.06 | 0.00 | 2.40 | 2.00 | 2008 |
| Change in the Length of Stay (in years) for Each Subsequent Sentence | | | | | | | | | | |
| Adult | 0.1839 | 0.1839 | 0.1839 | 0.1839 | 0.1839 | 0.1839 | 0.1839 | | | |
| Juvenile | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | | | |
| Age when a juvenile is first tried in adult court | | | | | | | | | | |
| Age when juvenile is tried as an adult | 16 | 16 | 16 | 18 | 18 | 18 | 18 | | | |

Change in the Length of Stay for Each Subsequent Sentence. In Washington, the sentence for a crime is based on the seriousness of the offense and the offender's criminal history. The Sentencing Guidelines Commission (SGC) publishes a grid showing the sentence by seriousness and the number of previous convictions. The sentence length for a given crime increases as criminal history increases.

To account for these lengthening sentences, we use the sentencing grid and the Institute's average length of stay data to create a new sentencing grid weighted for the frequency of conviction and the likelihood of prison. This enables us to estimate the effect of increasing trips through the criminal justice system on sentence length.

⁶⁷ Washington State Sentencing Guidelines Commission (personal communication, April 6, 2010).

We estimate this first, by determining the average length of stay for recidivists convicted of the following offense categories: murder, sex, robbery, assault, property, drug, and misdemeanor. We assume offenders who released from prison have at least three prior offenses and then determine the following:

- Likelihood of conviction.
- Likelihood of going to prison if convicted.
- Average length of stay (LOS).

Next, we determine what the offense seriousness level is upon the fourth conviction. We do this by matching the length of stay for the offense category with the seriousness level in the sentencing grid and with a sentence most similar to the length of stay. For example, the average length of stay in prison for murder (all offenses from manslaughter through first degree murder) is 13.4 years. This length of stay, with three prior offenses, is closest to the sentence at Seriousness Level XIII.

We then weight the sentences in the grid for the likelihood of recidivism in the offense categories and the likelihood of going to prison.

Finally, we create a single grid with increased average sentences by increased number of prior convictions. We plot this weighted average sentence by number of offenses. The result is a linear relationship; the slope indicates that each subsequent conviction increases the average prison sentence by an additional 0.1839 year. As of August 2010, we have not computed a similar procedure for juvenile repeat offenders sentenced to state institutions.

Age When a Juvenile Is First Tried in Adult Court. Under Washington's current laws, the age at which a youth is considered an adult varies for specific types of crimes. Exhibit D2.p contains information on the maximum age for juvenile court jurisdiction by type of crime. The actual determination of juvenile or adult court jurisdiction depends on several factors, in addition to a person's age and his or her crime. The model uses the information in Exhibit D2.p as representative of the typical decisions made pursuant to current Washington State law.

D2.2 Criminological Information for Different Populations

To estimate the long-run impacts of evidence-based programs on crime, the Institute combines program effect sizes with crime information from various populations in Washington State. To do this analysis, we calculate 15-year recidivism trends for an offender cohort; for non-offender populations, we calculate the probability of obtaining a conviction over the life-course (35 years).

Crime Parameters. As shown in Exhibit D2.q, we calculate the following information for each of the offender and non-offender populations:

- Conviction Rate. We estimate the cumulative conviction rate for felony and misdemeanor crime in Washington over the 15-year follow-up period. We compute the cumulative conviction rate using a fourth order polynomial density distribution. These conviction rates, by year, are used to calculate the unit change in crime as described in section C2 of this appendix.
- Crime Probability. For people who do commit crimes during the follow-up period, we calculate the probability of being convicted for a certain type of crime using a ranked order of seriousness. The mutually exclusive categories from most serious to least serious include: murder, sex, robbery, assault, property, drug, and misdemeanor.
- Trips Through the System. We calculate the total number of adjudications, defined as the number of "trips" through the criminal justice system, during the follow-up period. We also determine the average number of trips per offender during the follow-up period.
- Volume of Offenses. It is possible for offenders to have multiple offense convictions for each trip through the system. Thus, we also calculate the total number of offenses during the follow-up period, as well as the average number of offenses per adjudication. Adjudications and offenses are broken into the following categories: murder, sex, robbery, assault, property, drug, and misdemeanor.
- Timing. For those persons convicted, we compute a probability density distribution for each of the offender and non-offender populations using exponential, lognormal, polynomial (second order), or power distributions, which indicate when convictions are likely to happen over the follow-up period.

Exhibit D2.q

WSIPP Benefit-Cost Model: Version 1.2

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

Crime

Back to Main Model

Per Unit Costs | Resource Use | Offender Populations | Victimization

Select the type of population group to View/Modify

Adult Prison-General

Offender population name: Adult Prison-General

Number years follow-up: 15

Density distribution parameters

| Cumulative recidivism rate (conviction rate) | | Hazard rate (timing) | |
|--|---------|----------------------|--------|
| Parameter 1 | 0.261 | Distribution type | 2 |
| Parameter 2 | 0.1991 | Parameter 1 | 0.1279 |
| Parameter 3 | -0.0315 | Parameter 2 | -0.033 |
| Parameter 4 | 0.0022 | Parameter 3 | 0 |
| Parameter 5 | -0.0001 | Parameter 4 | 0 |

| | Murder | Felony Sex Crimes | Robbery | Aggravated Assault | Felony Property | Felony Drug | Misdemeanor | Total |
|---|--------|-------------------|---------|--------------------|-----------------|-------------|-------------|-------|
| Crime probability: most serious recidivism offense | 0.0132 | 0.0401 | 0.0759 | 0.2433 | 0.2521 | 0.1582 | 0.2172 | 1 |
| Trips: average number of adjudications through the system | 1.091 | 1.08 | 1.392 | 1.546 | 3.831 | 5.665 | 13.754 | |
| Offenses: average number of offenses per trip | 1.417 | 2.213 | 1.662 | 1.443 | 1.567 | 1.27 | 1.141 | |

Offender Populations. Recidivism is defined as any offense committed after release to the community, or after initial placement in the community, that results in a conviction in Washington State from adult or juvenile court.⁶⁸ In addition to the 15-year follow-up period, a one-year adjudication period is added to allow for court processing of any offenses that occur at the end of the follow-up period. Crime parameters are calculated using the Institute's criminal history database, which is a synthesis of information for offenders convicted in Washington State.⁶⁹

We collected recidivism data on five general offender populations who became "at-risk" for recidivism in the community during calendar year 1993. For adult offenders, we observe recidivism patterns for (1) offenders released from Department of Corrections' (DOC) facilities, and (2) offenders sentenced directly to DOC community supervision. For juvenile offenders, we observe recidivism patterns for (3) youth released from Juvenile Rehabilitation Administration (JRA) facilities, (4) youth sentenced to diversion through local-sanctioning courts, and (5) youth sentenced to detention/probation through local-sanctioning courts. We calculated separate crime distributions for each offender population.

We further break down the general offender populations into risk for reoffense categories. Risk for reoffense is calculated using criminal history data to determine offenders' probability of future reoffense, and grouped into low, moderate, and high risk categories.⁷⁰ Additionally, we created and analyzed adult and juvenile sex offender populations based on the most serious current offense of conviction prior to the 15-year follow-up period.

Non-Offender Population. To determine the impact of prevention programs on future crime, we calculate the probability of obtaining a conviction over the life-course for a birth cohort. We select felony and misdemeanor offenders from the Institute's criminal history database who were born in 1974 (n=78,517) to determine how many people were convicted at age 8, age 9, age 10, and so on. The 1974 birth cohort gives us the longest follow-up period (36 years) possible using Washington State criminal records data. Using Office of Financial Management state population data, we abstract the number of people living in Washington State, and born in 1974, for each of the follow-up years. For example, in 1994, there were 66,709 20-year-olds (1974 birth cohort) living in Washington. We estimate the average size of the 1974 cohort each follow-up year weighted by crime propensity. Future crime probability is adjusted as the life-course progresses.

Low Income Non-Offender Population. We also estimate criminological information for a low income population by adjusting the non-offender population described above using poverty and arrest data from the National Survey on Drug Use and Health.⁷¹ Specifically, we estimate for the low income population (1) a new base conviction rate over the life-course and (2) the probability of being convicted for a certain crime.

To do this, we use multivariate logistic regression analysis to determine the effect of poverty on crime with arrests as the dependent variable and poverty as the independent variable along with relevant control variables (See Exhibit D2.r). Poverty is measured as less than 200 percent of the federal poverty threshold. The coefficient from this model indicates that poverty is significantly related with a greater likelihood of crime ($b=.803$, $p<.0001$). We use the coefficient to adjust the base conviction rate over the life-course by calculating the odds ratio multiplied by the base conviction rate at any year over the life-course, divided by the odds ratio of the base conviction rate remaining over the life-course (for example, $\exp(.803) \cdot .33 / ((1-.33) + .33 \cdot \exp(.803))$).

We adjust the probability of being convicted for a certain type of crime by conducting individual multivariate regression analyses for arrests for a violent crime, arrests for a property crime, arrests for a drug crime, and arrests for other crime. We took the ratio of the odds ratios for each of those crime categories relative to the total poverty effect. We multiplied the ratio of odds ratios by the crime probability for the non-offender population, and normalized the crime probability to 1.

⁶⁸ Barnoski, R. (1997, December). *Standards for improving research effectiveness in adult and juvenile justice* (Document No. 97-12-1201). Olympia: Washington State Institute for Public Policy, pg. 2.

⁶⁹ Criminal history data are from the Washington State Administrative Office of the Courts and Department of Corrections.

⁷⁰ See Barnoski, R., & Drake, E. (2007, March). *Washington's Offender Accountability Act: Department of Corrections' static risk instrument* [Revised October, 2008] (Document No. 07-03-1201). Olympia: Washington State Institute for Public Policy. See also, Barnoski, R. (2004, March). *Assessing risk for re-offense: Validating the Washington State juvenile court assessment* (Document No. 04-03-1201). Olympia: Washington State Institute for Public Policy.

⁷¹ U. S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2010, November 16). *National Survey on Drug Use and Health, 2009* [Computer file]. ICPSR29621-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. doi:10.3886/ICPSR29621

Exhibit D2.r

| | Type of Arrest | | | | |
|------------------|-----------------------|----------------|-----------------|-------------|--------------|
| | Any | Violent | Property | Drug | Other |
| Intercept | -4.717 | -6.457 | -7.024 | -7.062 | -5.111 |
| Poverty | 0.803 | 1.013 | 1.126 | 0.630 | 0.653 |
| Male | 1.148 | 1.213 | 0.726 | 1.039 | 1.196 |
| Age 12-13 | -1.095 | -0.269 | 0.623 | 0.038 | -2.160 |
| Age 14-15 | 0.157 | 0.734 | 1.606 | 0.769 | -0.667 |
| Age 16-17 | 0.598 | 0.850 | 1.847 | 1.525 | -0.160 |
| Age 18-20 | 1.058 | 0.864 | 1.904 | 1.827 | 0.700 |
| Age 21-25 | 0.978 | 0.772 | 1.277 | 1.908 | 0.733 |
| Age 26-34 | 0.676 | 0.645 | 1.498 | 0.880 | 0.517 |
| Black | 0.462 | 0.653 | 0.286 | 0.512 | 0.321 |
| Native American | 1.008 | 1.613 | -0.168 | 0.601 | 0.815 |
| Pacific Islander | 0.161 | -0.253 | -0.666 | -0.444 | 0.443 |
| Asian | -1.615 | -3.029 | -2.317 | -1.766 | -1.235 |
| Hispanic | 0.052 | 0.299 | -0.202 | -0.496 | 0.094 |
| Married | -1.019 | -1.172 | -1.027 | -1.291 | -0.990 |
| Model Fit | 0.750 | 0.752 | 0.734 | 0.778 | 0.746 |

All variables were statistically significant for all models at $p < .001$.

D2.3 Estimates of Victimitizations Per Conviction

Once a person is convicted for a criminal offense, sentencing policies and practices in Washington affect the use of different local and state criminal justice resources. Exhibit D2.s is a screen shot from the Institute's benefit-cost model which displays how criminal justice resources in Washington State are used in response to crime. Yellow boxes contain inputs entered by the Institute while blue boxes contain calculations. Inputs in Exhibit D2.s are described below.

Exhibit D2.s

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

General | Economic | **Crime** | Education | Child Welfare | Substance Use | Health Care | Mental Health | Public Asst | Housing | Teen Birth | Outcomes & Links

Crime

Back to Main Model

Per Unit Costs | Resource Use | Offender Populations | **Victimization** | Program Participation

| | Murder | Rape | Robbery | Aggravated Assault | Burglary | Theft | Motor Vehicle Theft | Year of Data |
|--|--------|-------|---------|--------------------|----------|--------|---------------------|--------------|
| Number of statewide crimes reported to police | 191 | 2664 | 6345 | 12451 | 52664 | 166214 | 28715 | 2008 |
| Multiplicative adjustment to align with felonies | 1 | 2.41 | 1 | 1 | 1 | 0.235 | 1 | |
| Number of statewide adjusted crimes reported to police | 191 | 6420 | 6345 | 12451 | 52664 | 39060 | 28715 | |
| Percent of crime reported to police | 1 | 0.307 | 0.656 | 0.572 | 0.501 | 0.685 | 0.853 | 2007 |
| Statewide estimated felony-type crimes | 191 | 20912 | 9672 | 21767 | 105118 | 57022 | 33664 | |
| Statewide number of convictions, adult and juvenile | 240 | 1680 | 813 | 4437 | 11875 | 10917 | | 2008 |
| Statewide number of counts, adult and juvenile | 328 | 3338 | 1277 | 7223 | 24627 | | | |
| Average number of offenders per victim | 1 | 1 | 1 | 1 | 1 | | | |
| Statewide estimated felony-type crimes | 191 | 20912 | 9672 | 21767 | 195804 | | | |
| Percent of other crimes per conviction | 0.64 | 0.2 | 0.2 | 0.2 | 0.2 | | | |
| Estimated victimizations per convicted offender | 1 | 4.08 | 3.64 | 2.28 | 4.96 | | | |
| Variance in ratios of victimizations per convicted offender Low Percent: -0.2 High Percent: 0.2 | | | | | | | | |
| Statewide number of arrests, adult and juvenile | 148 | 1918 | 1892 | 5456 | 35819 | 28261 | | 2008 |
| Maximum number of arrests per conviction | 0.62 | 1.14 | 2.33 | 1.23 | 3.02 | 2.59 | | |
| Percent of other arrests attributed to a conviction | 0 | 0 | 0 | 0 | 0 | 0 | | |
| Estimated number of arrests per conviction | 1 | 1 | 1 | 1 | 1 | 1 | | |

Number of statewide crimes reported to the police. Uniform Crime Report (UCR) data for all policing agencies are obtained from the Washington Association of Sheriffs and Police Chiefs. We adjust the data to account for non-reporting agencies. The data are then aggregated to statewide annual data.

Multiplicative adjustment to align UCR data with Washington felonies. Two of the UCR-reported crime categories, rape and felony theft, do not align with felony conviction data as defined by the Revised Code of Washington. Thus, we apply a multiplicative adjustment factor to align reported crimes with felony convictions.

Rape, as defined by the UCR, does not include other sexual assaults, sexual offenses with male victims, or victims under the age of 12. We adjust UCR reported rapes using National Crime Victimization Survey data to estimate male victims⁷² and other sexual assaults.⁷³ Data from the National Incident Based Reporting System are used to adjust for the percentage of all sex offenses where victims are under age 12.⁷⁴

⁷² Bureau of Justice Statistics. (2008, August). *Criminal victimization in the United States, 2006 statistical tables: National Crime Victimization Survey* (Document No. NCJ 223436), Washington, DC: United States Department of Justice, Author, Table 2.

⁷³ Ibid., Table 1.

⁷⁴ Snyder, H. N. (2000, July). *Sexual assault of young children as reported to law enforcement: Victim, incident, and offender characteristics* (Document No. NCJ 182990). Washington, DC: United States Department of Justice, Bureau of Justice Statistics.

Theft is adjusted to include only thefts valued at \$750 or more, the cutoff for a felony theft, as defined by the Revised Code of Washington. We use National Crime Victimization Survey data of thefts reported to the police to estimate this figure.⁷⁵

Percentage of crimes reported to the police. We adjust our victimization estimates to include crimes not reported to the police using reporting rate data obtained from the National Crime Victimization Survey.⁷⁶

Statewide number of convictions, adult and juvenile. Adult and juvenile felony conviction data are obtained from the Administrative Office of the Courts.⁷⁷

Average number of offenders per victim. Many victimizations are committed by groups of offenders, thus we estimate the average number of offenders per victimization using data from the National Incident Based Reporting System (NIBRS).⁷⁸ We use the offender sequence number in the NIBRS data, which indicates the number of offenders for each incident, and we determine the average number of offenders for each broad offense category.

Percentage of other crimes per conviction. In order to estimate the total number of crimes per convicted offender, we apply a multiplicative factor to adjust for the likely possibility that there are multiple victimizations per conviction. To our knowledge, no research exists to date that indicates the appropriate value. Thus, we simply apply an estimate of 20 percent. A value of zero would imply one victimization per conviction, while a value of one would imply all crimes are attributed to those offenders convicted.

Statewide number of arrests, adult and juvenile. Arrest data were obtained from the Washington Association of Sheriffs and Police Chiefs. We adjust the data to account for non-reporting agencies. The data are then aggregated to statewide annual data.

Percentage of other arrests attributed to a conviction. There is a provision in the model to account for all other arrests attributed to a conviction; however, we do not currently use this information.

D2.4 Procedures to Estimate Criminal Justice System and Victimization Events

In this section of the technical appendix, we describe how the inputs from the previous sections are used to calculate victimizations and costs avoided. In some instances, we also count the quantity of criminal justice events, such as prison beds, avoided.

Criminal Justice System Resources.

For each criminal justice resource, as seen in Exhibit D2.a, we estimate costs avoided using the following equation:

$$(3) CjsResource\$_{ry} = \sum_{c=1}^C \sum_{t=1}^{ceiling(T_c)} \sum_{f=1}^F [CjsEvent_{yctf} \times CrimePr_c \times CjsResourcePrW_{rc} \times TripPr_{ct} \times TimetoRecid_f \times RelRisk_y \times CjsResourceCost_{rc}] \times RecidRate$$

We also count Average Daily Population prison beds avoided. We do this using equation 3 above however; we do not multiply by the $CjsResourceCost_{rc}$.

⁷⁵ Bureau of Justice Statistics, 2008, Table 100.

⁷⁶ Ibid.

⁷⁷ Washington State Administrative Office of the Courts, Superior Court Annual Tables, available from <http://www.courts.wa.gov/caseload/?fa=caseload.showIndex&level=s&freq=a>

⁷⁸ U. S. Department of Justice. Federal Bureau of Investigation. *National incident-based reporting system, 2008* [Computer file]. ICPSR27647-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-03. doi:10.3886/ICPSR27647.

Variable Definitions. Below are definitions and calculations for the variables used in Equation 3.

C – The number of crime types modeled, ranked from most serious crime category to least serious. For example, we use seven crime types ranked in the following order: murder, sex offenses, robbery, aggravated assault, property, drug, and misdemeanors.

F – The number of years in the recidivism follow-up.

Y – The at-risk year following treatment.

T – The number of trips (adjudications) through the system rounded up. For example, prison offenders, whose most serious reoffense is a sex offense, have an average number of 1.08 trips in a 15-year follow-up period. Thus, the total possible number of trips through the system is two with the probability of the second trip being less than .08. See also *TripPr_{ct}*.

CjsEvent_{yctf} – Variable indicating if and when a criminal justice resource is used and, if so, how much of the resource is used during the at-risk year. Criminal justice resources are shown in Exhibit D2.a. The Visual Basic Programming language for *CjsEvent_{yctf}* is shown in Exhibit D2.s.

CrimePr_c – Among those who re-offend, the probability that the most serious offense occurring during the follow-up period is of type *c*. The data for populations are show in Exhibit D2.q.

CjsResourcePrW_{rc} – The probability that a criminal justice resource will be used for a specific type of crime. See Exhibit D2.1p. For example, not all offenders who are convicted of a crime will necessarily receive a prison sentence. The *CjsResourcePrW_{rc}* for police and courts is 1.

TripPr_{ct} – The probability that a trip, a criminal justice event resulting in an adjudication during the follow-up period, occurs for crime *c* for trip *t* as show in Exhibit D2.q. The probability of a trip occurring is 1. Once a whole trip has been used, then we use the remaining probability of the trip. For example, prison offenders whose most serious reoffense is a sex offense have an average number of 1.08 trips in a 15-year follow-up period. Thus, there is a probability of 1 trip occurring and a probability of .08 remaining trips.

TripSpaces_c – The number of years in the follow-up period divided by the number of *Trips_c*. This estimate enables us to distribute the total number of adjudications over the 15-year period.

TimetoRecid_t – Among those who re-offend during the recidivism follow-up period *f*, the probability that the recidivism event happens in year *f*. The sum of *TimetoRecid_f* = 1.0

RelRisk_y – The change in the relative risk in crime outcomes in year *y*. Equation 4 shows how we calculate *RelRisk_y*.

$$(4) RelRisk_y = \left[\frac{\left(\frac{e^{ES \times 1.65} \times RecidRate}{(1 - RecidRate + RecidRate \times e^{(ES \times 1.65)})} \right)}{RecidRate - 1} \right]$$

ES – The estimated effect size on crime outcomes for the program. The value is computed as a standardized mean difference effect size, approximated for dichotomous outcomes with the Dcox transformation.

CjsResourceCost_{rc} – The per unit marginal costs of each criminal justice resource as estimated in section D.2 of this appendix and as shown in Exhibit D2.a

RecidRate – The percentage of offenders who have a Washington state court legal action during the recidivism follow-up period *F* for that specific offender population as shown in Exhibit D2.q. Different recidivism base rates are used depending on the specific population that receives a given program. See Exhibit D2.q.

Exhibit D2.rt

Visual Basic Programming Code Used to Calculate $CjsEvent_{yctf}$

```

RowCount = 0
For c = 1 To CrimeTypes
  For t = 1 To TripsCeiling(c, 1)
    If t <= Trips(c, 1) Then
      TripMultiplier = 1
    Else
      TripMultiplier = Trips(c, 1) - Int(Trips(c, 1))
    End If
    AgeTemp = age + (t - 1) * TripSpaces(c, 1)

    For f = 1 To FollowUpYears
      RowCount = RowCount + 1
      If AgeTemp < AgeofAdultCJS(c, 1) Then GoTo SkipAdult

      For y = 1 To MaxAtRiskYears
        If (f + ((t - 1) * TripSpaces(c, 1))) > y Then
          CjsEvent(RowCount, y) = 0
        ElseIf Int(CjsResourceLength) + (f + ((t - 1) * TripSpaces(c, 1))) = y Then
          CjsEvent(RowCount, y) = CjsResourceLength - Int(CjsResourceLength)
        ElseIf y > CjsResourceLength + (f + ((t - 1) * TripSpaces(c, 1))) Then
          CjsEvent(RowCount, y) = 0
        Else
          CjsEvent(RowCount, y) = 1
        End If

        CjsResourceAvoided(1, y) = CjsResourceAvoided (1, y) _
          + CjsEvent(RowCount, y) _
          * CrimeProbCjs(c, 1) _
          * CjsResourceProb(c, 1) _
          * TimeToRecid(f, 1) _
          * TripMultiplier _
          * RelativeRisk(f, 1) _
          * RecidRate _
          * CjsResourcePerUnitCost(c, 1) _
          * (1 + CjsResourcePerUnitCost Esc) ^ (y - 1)
      Next y
    SkipAdult:
      AgeTemp = AgeTemp + 1
    Next f
  Next t
Next c
For y = 1 To MaxAtRiskYears
  CjsResourceAvoided (y, 1) = CjsResourceAvoidedSum(1, y)
Next y

```

Victimizations Avoided

Using information from Exhibit D2.q, we estimate the number of victimizations avoided and victimization costs avoided using the following equation:

$$(5) \text{Victim}\$_{ry} = \sum_{c=1}^C \sum_{t=1}^{\text{ceiling}(\text{Trips}_c)} \sum_{f=1}^F [\text{VictimEvent}_{yctf} \times \text{VictimVolume}_{yctf} \times \text{CrimePr}_c \times \text{TripPr}_{ct} \times_f \text{RelRisk}_y \times \text{VictimCost}_{rc}] \times \text{RecidRate}$$

Variable Definitions. Below are definitions and calculations for the variables used in Equation 5 unless otherwise defined in the aforementioned section, criminal justice system resource variable definitions.

VictimVolume_{ctf}. The volume of victimizations is estimated using a three-step process. First, we estimate the number of victimizations avoided for the most serious offense in the follow-up period. Second, since there are usually other offenses adjudicated at the time of the most serious offense, we calculate the additional offenses and related victimizations. Finally, we calculate the number of victimizations avoided for the trips through the criminal justice system during the remainder of the follow-up period.

F – The number of years in the recidivism follow-up time trips ceiling for that offense type. For example, prison offenders whose most serious reoffense is a sex offense have an average number of 1.08 trips in a 15-year follow-up period. Thus, the ceiling of the total number of trips that need to be modeled are 2.

$$(6) \text{VictimVolume}_{cf} = \sum_{c=1}^C \sum_{v=c}^C \sum_{f=1}^F \frac{(\text{MostSeriousTripVic}_c + \text{AddVicsMostSeriousTrip}_c + \text{RemainingTrips}_c)}{\text{Trips}_c}$$

Equations 7, 8, and 9 show our calculations for each component of *VictimVolume_{yctf}*. In the following equations, when *c* equals *v*, we estimate the most serious offense using the following formulas. Otherwise, *c*, the most serious crime, is equal to zero.

$$(7) \text{MostSeriousTripVic}_c = 1 \times \text{VicsPerConvictedOffender}$$

$$(8) \text{AddVicsMostSeriousTrip}_c = \text{OffensesPerTrip}_c \times \text{VicsPerConvictedOffender}_v \times \left(\frac{\text{CrimePr}_v}{\sum_{c=1}^C \text{CrimePr}_c} \right)$$

$$(9) \text{RemainingTrips}_c = (\text{Trips}_c - 1) \times \text{OffensesPerTrip}_c \times \text{VicsPerConvictedOffender}_c \times \left(\frac{\text{CrimePr}_v}{\sum_{c=1}^C \text{CrimePr}_c} \right)$$

VictimEvent_{yctf}. A dichotomous variable indicating if a victimization event has occurred during the at-risk year. Victimizations are shown in Exhibit D2.s. The Visual Basic Programming language for *VictimEvent_{yctf}* is shown in Exhibit D2.u.

VictimCost_{rc}. The per unit cost of crime to victims as estimated in section D.2 of this appendix and as shown in Exhibit D2.a.

Exhibit D2.u

Visual Basic Programming Code Used to Calculate $VictimEvent_{yctf}$

```
For v = 1 To CrimeTypes
    RowCount = 0
    For c = 1 To CrimeTypes
        For t = 1 To TripsCeiling(c, 1)
            If t <= Trips(c, 1) Then
                TripMultiplier = 1
            Else
                TripMultiplier = Trips(c, 1) - Int(Trips(c, 1))
            End If
            For f = 1 To FollowUpYears
                RowCount = RowCount + 1
                For y = 1 To MaxAtRiskYears
                    If f + (t - 1) * TripSpaces(c, 1) = y Then
                        VictimEvent(RowCount, y) = 1
                    Else
                        VictimEvent (RowCount, y) = 0
                    End If
                    AvoidedVictims(v, y) = AvoidedVictims(v, y) _
                        + VictimEvent (RowCount, y) _
                        * CrimeProbCjs(c, 1) _
                        * TimeToRecid(f, 1) _
                        * TripMultiplier _
                        * RelativeRisk(f, 1) _
                        * RecidRate _
                        * VictimVolume(RowCount, v)
                Next y
            Next f
        Next t
    Next c
Next v
```

D3. Valuation of Child Abuse and Neglect Outcomes

The Institute's benefit-cost model contains procedures to estimate the monetary value of changes in the occurrence of child abuse and neglect (CAN), as well as the monetary value of changes in out-of-home placement (OoHP) in the child welfare system. This section of the Technical Appendix describes the Institute's current procedures to estimate the monetary benefits of program-induced changes in CAN and OoHP.

In general, analysts have constructed two types of studies to estimate the costs of CAN: "prevalence-based" studies and "incidence-based" studies. Prevalence costing studies look backward and ask: How much does CAN cost society today, given all current and past CAN among people alive in a state or country?⁷⁹ Incidence costing studies look forward and ask: How much benefit could be obtained in the future if CAN was reduced? Both approaches use some of the same information, but assemble it different ways. Incidence-based studies are more useful for estimating the expected future benefits and costs of policy choices; the Institute's model uses an incidence-based approach.

This component of the Institute's benefit-cost model is designed to ascertain whether or not there are effective, economically attractive policy options that can reduce CAN and OoHP if implemented well. The Institute's model includes estimates for the value of reducing a substantiated child abuse and neglect (CAN) case, from the perspective of the victim, and to society at large. In addition, we estimate the value of avoiding out-of-home placements in foster care, from the perspective of the taxpayer. The direct benefits are derived by calculating the costs that are incurred with the incidence of a child abuse and neglect case, or an occurrence of placement out-of-home.

CAN costs are a function of three principal components: the expected value of public costs associated with a substantiated CAN case (e.g., child welfare system and court costs), and an estimate of the medical, mental health, and quality of life costs associated with the victim of CAN. Other long-term costs that are causally linked to the incidence of CAN are discussed in Appendix E. OoHP costs are derived from the expected value public costs of an out-of-home placement, conditional on that placement occurring. As the costs for OoHP are most often a function of CAN-related participation in the child welfare system, we most frequently refer to the "CAN model" when describing our computations below.

Limitations of Our Methods for Valuing Reductions in CAN and OoHP

In the current benefit-cost model, we do not estimate the benefits of reducing CAN to the *children* of CAN victims. Our model is presently limited to effects on the two generations of CAN prevention or intervention program participants: the parent and the child (potential victim). Some research has demonstrated that CAN victims are more likely to perpetrate abuse or neglect on their own children;⁸⁰ we are unable to monetize those effects at this time. Second, there is a higher risk of death among CAN victims compared to other children. In our model, we do not monetize these deaths explicitly, but rather through our valuation of victim costs. Because we do not model death from CAN explicitly, we do not compute benefits derived from death adjusted life years (DALY) or the value of a statistical life. For victimization costs that include the probability of death from CAN, we use estimates from Miller, Fisher, and Cohen, which are discussed in some detail below.⁸¹

Finally, our model describes the direct result of a reduction in CAN by calculating the reduced public spending by the agencies that process CAN cases and a reduction in CAN victimization costs. In addition to these direct benefits, however, a reduction in CAN can also be expected to have an indirect causal linkage to several other outcomes monetized in this benefit-cost analysis. For example, there is credible research showing a causal link between the incidence of CAN and subsequent criminality of the victimized youth when he or she is older. Thus, when a prevention program is able to demonstrate an effect on the rate of child abuse and neglect, it is important to measure both the direct and indirect benefits that can be expected as a result.

For a complete description of the links between CAN and other outcomes later in life, see Appendix E.

⁷⁹ See for example, Wang, C.-T., & Holton, J. (2007, September). *Total estimated cost of child abuse and neglect in the United States*. Chicago: Prevent Child Abuse America. Retrieved June 30, 2011 from:

http://www.preventchildabuse.org/about_us/media_releases/pcaa_pew_economic_impact_study_final.pdf

⁸⁰ Whipple, E. E., & Webster-Stratton, C. (1991). The role of parental stress in physically abusive families. *Child Abuse & Neglect*, 15(3), 279-291; Hunter, R. S., Kilstrom, N., Kraybill, E. N., & Loda, F. (1978). Antecedents of child abuse and neglect in premature infants: A prospective study in a newborn intensive care unit. *Pediatrics*, 61(4), 629-635; Kim, J. (2009). Type-specific intergenerational transmission of neglectful and physically abusive parenting behaviors among young parents. *Children and Youth Services Review*, 31(7), 761-767; Belsky, J. (1993). Etiology of child maltreatment: A developmental-ecological analysis. *Psychological Bulletin*, 114(3), 413-434.

⁸¹ Miller, T. R., Fisher, D. A., & Cohen, M. A. (2001). Costs of juvenile violence: Policy implications. *Pediatrics*, 107(1). DOI: 10.1542/peds.107.1.e3

D3.1 Input Screens for CAN Parameters

The CAN model is driven with a set of parameters describing various aspects of CAN epidemiology, participation in the child welfare system, and linked relationships with other outcomes. These input parameters are shown in the following three screen shots. In addition, there are several other input parameters used in the CAN model that are general to the Institute’s overall benefit-cost model; these are discussed elsewhere in this Appendix. In the following sections, the sources for the parameters and the computational routines are described.

Exhibits D3.1a, D3.1b, and D3.1c display the parameters for the analysis of child abuse and neglect and out-of-home placement in the child welfare system. Each is described in detail below.

Exhibit D3.1a

The screenshot displays the 'Child Welfare' input screen within the WSIPP Benefit-Cost Model (Version 1.1). The interface is organized into several sections:

- Navigation:** A vertical menu on the left lists categories: General, Economic, Crime, Education, Child Welfare (selected), Substance Use, Health Care, Mental Health, Public Asst, Housing, Teen Birth, and Outcomes & Links. The top menu bar includes 'Enter Sector Inputs', 'Enter Program Inputs', and 'Run Models & View Reports'.
- Child Welfare Section:** A sub-menu at the top left of the main area includes 'Child Welfare Sytem', 'Victimization Costs', and 'Prevalence Rates'. A 'Back to Main Model' button is also present.
- Base Likelihood of Out-of-Home Placement:** This panel contains input fields for:
 - All Children: 0.06
 - Children at Imminent Risk of Removal: 0.75
 - Children with Serious Emotional Disturbance: 0.43
 - Decay rate for timing of CAN costs, following an incident of CAN:
 - System costs: -0.51
 - Victim costs: -0.1
- Child Abuse and Neglect: Prevalence in the General Population:** This panel features a table showing the 'General Population Rate of First Substantiation by Age (Cumulative)'.

| Age | Proportion |
|-----|------------|
| 0 | 0.0206 |
| 1 | 0.0292 |
| 2 | 0.0372 |
| 3 | 0.0447 |
| 4 | 0.0515 |
| 5 | 0.0583 |
| 6 | 0.0650 |
| 7 | 0.0717 |
| 8 | 0.0772 |
| 9 | 0.0827 |
| 10 | 0.0881 |
| 11 | 0.0935 |
| 12 | 0.0986 |
| 13 | 0.1036 |
| 14 | 0.1086 |
| 15 | 0.1136 |
| 16 | 0.1169 |
| 17 | 0.1203 |

 Below the table, the 'Proportion' is set to 0.0206, and the 'Odds Ratio for Low-SES population' is 2.175.
- Child Abuse and Neglect: Recurrence for Maltreated Children:** This panel shows a table for the 'Indicated Population Recurrent Substantiation by Follow-up Year (Cumulative)'.

| Follow-up Year | Proportion |
|----------------|------------|
| 1 | 0.2362 |
| 2 | 0.3009 |
| 3 | 0.3376 |
| 4 | 0.3637 |
| 5 | 0.3839 |
| 6 | 0.4004 |
| 7 | 0.4143 |
| 8 | 0.4264 |
| 9 | 0.4370 |
| 10 | 0.4466 |
| 11 | 0.4552 |
| 12 | 0.4631 |
| 13 | 0.4703 |
| 14 | 0.4770 |
| 15 | 0.4833 |
| 16 | 0.4891 |
| 17 | 0.4946 |
| 18 | 0.4998 |

 The 'Proportion' field is currently empty.

Exhibit D3.1b

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

Child Welfare

Back to Main Model

Child Welfare System | Victimization Costs | Prevalence Rates

Child Welfare System Costs for a Child Abuse and Neglect Case

Child Protective Services (CPS)- Related Caseloads and Costs

| | Annual Number of Children | Year of data | Probability of getting this service | Dollars Per Child | Year of Dollar Estimates | Dollar Total |
|---|---------------------------|--------------|-------------------------------------|-------------------|--------------------------|--------------|
| Referrals (children) accepted for CPS investigation | 42800 | 2008 | 1 | 618 | 2008 | 618 |
| Police Involvement | 5093 | 2008 | 0.119 | 670 | 2009 | 80 |
| Juvenile Court Involvement (Dependency cases) | 3883 | 2009 | 0.091 | 3547 | 2009 | 322 |

Child Welfare Services (CWS)- Related Caseloads and Costs

| | | | | | | |
|--|--------|------|-------|---------|------|------|
| Percentage of Placements due to Child Maltreatment | 0.7527 | | | | | |
| Protective Custody (new placements) | 7500 | 2008 | 0.132 | 24568.5 | 2008 | 3243 |
| In-Home Services (not out-of-home placement) | 0 | 2008 | 0 | 0 | 2008 | 0 |
| Adoption | 733 | 2008 | 0.017 | 78656.6 | 2008 | 1337 |
| Juvenile Court Involvement (Termination cases) | 1934 | 2009 | 0.045 | 2640 | 2009 | 119 |

Expected present value cost of an accepted CPS case: 5719

Expected present value cost of an out-of-home placement, conditional on an out-of-home placement: 32936.64

Expected length of stay in placement (in years), conditional on an out-of-home placement: 1.63

Expected cost of out-of-home placement for a youth with serious emotional disorder, conditional on an out-of-home placement: 7207

Year of data: 2010

Exhibit D3.1c

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

General | Economic | Crime | Education | **Child Welfare** | Substance Use | Health Care | Mental Health | Public Asst | Housing | Teen Birth | Outcomes & Links

Child Welfare

Back to Main Model

Child Welfare Sytem | Victimization Costs | Prevalence Rates

Child Welfare Victim Costs for a Child Abuse and Neglect Case

| | Direct Costs | Indirect Costs | Year of dollars |
|------------------------------------|--------------|----------------|-----------------|
| Estimated Lifetime Cost Per Victim | 1901 | 22948 | 1993 |
| Proportion paid by Victim | 0.5 | 1 | |
| Proportion paid by Taxpayer | 0.5 | 0 | |
| Proportion paid by Other | 0 | 0 | |

D3.2 CAN Parameters: Prevalence for Prevention and Intervention Programs, Timing of Costs

The Institute’s CAN model begins by analyzing the epidemiology of CAN to produce estimates of the cumulative likelihood of experiencing child abuse or neglect. An estimate of the cumulative prevalence of CAN is central to the benefit-cost model because it becomes the “base rate” of CAN to which program or policy effect sizes are applied to calculate the change in the number of avoided CAN “units” caused by the program, over the lifetime following treatment.

Exhibit D3.1a displays the input form for the cumulative prevalence of CAN, from age 1 to age 100.

To compute the estimated probability of being a victim of child abuse or neglect, we use data from the National Child Abuse and Neglect Data System to calculate the one-year prevalence of child victims by age group.⁸² In any given year, some of these cases are repeat cases from previous maltreatment episodes. We estimate this number by subtracting the proportion of first-time victims⁸³ from one. Using these two parameters to calculate the annual probability of a new substantiated child abuse or neglect case for a child from age zero to age 17, the implied lifetime prevalence rate of child abuse or neglect for the general population of children is estimated to be 12.0 percent. The cumulative prevalence for CAN by age, after repeat cases are accounted for, is displayed in Exhibit D3.1a.

To test the reasonableness of this estimate, we use a second approach to calculate the lifetime prevalence. We gathered other research studies that have examined this question with longitudinal cohort data. Exhibit D3.2a summarizes these estimates. The studies measured child abuse and neglect with different definitions, for different populations, and at different times. Ignoring these variations, a simple weighted average of the studies produces an estimate of 10.6 percent lifetime prevalence of child abuse, slightly lower than, but similar to the estimate described in the first method above.

Some of the populations that are the focus of prevention and early intervention programs are not the general population but are, instead, from higher risk populations, often from lower socio-economic status. For the model, we estimate a parameter for this (an odds ratio applied to the annual prevalence rate for the general population) by taking a weighted average of the results of five studies that have examined this question with control groups (see Exhibit D3.2b).⁸⁴

**Exhibit D3.2a
Lifetime Prevalence Estimates of Child Abuse and Neglect**

| Study | Number in study with abuse | Total number in sample | Percentage with child abuse or neglect | Notes |
|--------------------------------|----------------------------|------------------------|--|---|
| Total | 3,765 | 35,650 | 10.6% | Weighted average of studies listed |
| Eckenrode et al., 1993 | 1,239 | 8,569 | 14.5% | General pop, NY, substantiated cases |
| Stouthamer-Loeber et al., 2001 | 52 | 506 | 10.3% | Inner city pop, Pittsburg, substantiated |
| Zingraff et al., 1993 | 10 | 387 | 2.6% | School sample, Mecklenburg, NC |
| Thornberry et al., 2001 | 213 | 1,000 | 21.3% | Rochester, NY, substantiated cases |
| Reynolds et al., 2003 | 69 | 595 | 11.6% | Chicago higher risk sample, CPS control |
| MacMillan et al., 1997 | 1,461 | 9,953 | 14.7% | General pop, Ontario, severe, self report |
| Brown et al., 1998 | 46 | 644 | 7.1% | General pop, non SES |
| Kelleher et al., 1994 | 378 | 11,662 | 3.2% | Five urban sites |
| Dodge et al., 1990 | 46 | 304 | 15.1% | General pop, physical abuse |
| Finkelhor et al., 2003 | 252 | 2,030 | 12.4% | One year rate |

⁸³ Administration on Children, Youth and Families, (2009) *Child Maltreatment 2009* Table 3-10. Retrieved June 30, 2011, from <http://www.acf.hhs.gov/programs/cb/pubs/cm09/cm09.pdf>.

⁸³ *Ibid.*, Table 3-8.

⁸⁴ Lealman, G. T., Phillips, J. M., Haigh, D., Stone, J., & Ord-Smith, C. (1983). Prediction and prevention of child abuse—An empty hope? *The Lancet*, 321(8339), 1423-1424; Murphey, D. A., & Braner, M. (2000). Linking child maltreatment retrospectively to birth and home visit records: An initial examination. *Child Welfare*, 79(6), 711-728; Kotch, J. B., Browne, D. D., Dufort, V., Winsor, J., & Catellier, D. (1999). Predicting child maltreatment in the first 4 years of life from characteristics assessed in the neonatal period. *Child Abuse and Neglect*, 23(4), 305-319; Hussey, J. M., Chang, J. J., & Kotch, J. B. (2006). Child maltreatment in the United States: Prevalence, risk factors, and adolescent health consequences. *Pediatrics*, 118(3), 933-942; Brown, J., Cohen, P., Johnson, J.G., & Salzinger, S. (1998). A longitudinal analysis of risk factors for child maltreatment: Findings of a 17-year prospective study of officially recorded and self-reported child abuse and neglect. *Child Abuse and Neglect*, 22(11), 1065-1078.

Exhibit D3.2b
Odds Ratios for Child Abuse and Neglect: High Risk Populations

| Study | Study n | Odds ratio | High risk population |
|------------------------|---------------|--------------|--|
| Total | 43,707 | 2.175 | (weighted average) |
| Lealman et al., 1983 | 2,802 | 3.72 | Mothers under 20 OR with late prenatal care OR unmarried |
| Murphey & Braner, 2000 | 29,291 | 2.45 | Teen mothers OR eligible for medicaid |
| Kotch et al., 1999 | 708 | 1.36 | Receiving income support |
| Hussey et al., 2006 | 10,262 | 1.06 | Income less than \$15,000 |
| Brown, 1998 | 644 | 1.44 | Low income |

For children already in the child welfare system, we also estimate the likelihood of recurrence of abuse or neglect. The results of this analysis are displayed in Exhibit D3.1a; we combined the results of two studies that examined the recurrence of substantiated maltreatment between 1.5 and 5 years from the first substantiation.⁸⁵ Using data presented in these studies, we analyzed the proportion of children who had experienced a re-recurrence of abuse or neglect, from one month out to five years. We then plotted a logarithmic curve to predict the likelihood of a recurrence from 5 to 17 years after the initial incident.

Exhibit D3.1a also displays the base rates of out-of-home placement for various populations. For the general population, we calculated a lifetime probability of 6 percent based on an Institute analysis of Washington state child welfare data.⁸⁶ For the population of children already in the child welfare system deemed at “imminent risk” of placement, an Institute analysis⁸⁷ determined the risk of out-of-home placement for these children was much higher (75 percent), so our base rate of placement for programs that serve children at imminent risk is set at 75 percent. The third rate in Exhibit D3.1a shows the likelihood of out-of-home placement for children with serious emotional disturbance (SED). These children are sometimes placed in intensive foster care, or in the hospital for psychiatric treatment.⁸⁸

The final inputs in Exhibit D3.1a are the parameters that allow us to estimate the timing of costs incurred within the child welfare system. We have two rates of decay, one for costs within the child welfare system, and one for costs to the victim.

Within the system, costs for a case of child abuse or neglect do not occur all at once, but rather linger over time. Costs like an investigation, initial services to a family, dependency court, and so forth, occur early in the case, but child welfare services and out-of-home placements may continue for a number of years. From our data in Exhibit D3.1b, we estimated the amount of system-related costs we would expect to be incurred within the first two years of a typical CAN case (76 percent). Using that figure, we calculated a rate of “decay,” such that for each year after the beginning of a case, the amount of cost decayed by -.51. That means, in the first year, 51 percent of the total expected costs were incurred; by the end of the second year, 76 percent had been incurred; 88 percent by the end of the third year; and so on. This “decay” continues for a maximum of 17 years, as child welfare system costs for out-of-home placement, courts, and child welfare services, etc., typically do not continue past the age of 17.

We also estimated the amount of victim-related costs over time, expecting that these costs may linger much longer than system-related cost. Our estimated rate of decay for these costs was -.10, which means that, relative to system costs, we expect victim costs of mental health and quality of life to be spread over a greater number of years.

Estimated child welfare system costs are displayed in Exhibit D3.1b. The table below provides the sources for these figures, in some cases derived from Washington State data, and in other cases estimated from national data. We multiply the probability of receiving each service by the per-child cost to calculate an expected value cost for each accepted referral.

⁸⁵ Fluke, J. D., Yuan, Y. Y., & Edwards, M. (1999). Recurrence of maltreatment: An application of the National Child Abuse and Neglect Data System (NCANDS). *Child Abuse & Neglect*, 23(7), 633-650; DePanfilis, D., & Zuravin, S. J. (1999). Epidemiology of child maltreatment recurrences. *The Social Service Review*, 73(2), 218-223.

⁸⁶ Using data from DSHS CAMIS child placement data for fiscal year 2001, we counted the total number of unduplicated children in out-of-home placements. Of these children, we examined their entire placement history in Washington (back to 1993, the first year for which we have data), and determined the number with at least one prior placement. We found that of the 7,695 placed in FY 2001, 2,182 (28.0 percent) had a prior placement. Using these two parameters to calculate the annual probability of a new out-of-home placement for a child from age zero to age 17, the implied lifetime prevalence rate of out-of-home placement for the general population of children is estimated to be 6.0 percent.

⁸⁷ Institute analysis of evaluations of the Homebuilders® model of intensive family preservation services, which serve youth at “imminent risk” of placement. For children in the comparison groups of these evaluations, approximately 75 percent were indeed removed from home, after being deemed at “imminent risk.”

⁸⁸ We calculated the rate of 43 percent from five studies of Multisystemic Therapy for children with SED; these are the rates of placement for the comparison groups in those studies.

In addition, we calculate the expected cost of an out-of-home placement, conditional on a child being placed out-of-home. Therefore, the expected value cost of the *average* child who is the subject of an accepted CPS referral is \$5,719 in 2010 dollars. For a child who gets placed out-of-home, that cost is \$32,937.

Exhibit D3.2c
The Estimated Average Public Cost of a Child Protective Service Case Accepted for Investigation, State of Washington

| | Number of Children (1) | Probability of Receiving This Service (2) | Per-Child Cost (3) | Year of Dollar Estimates (4) | Expected Cost per Accepted Case (5) |
|--|---------------------------|--|---------------------------|---------------------------------|--|
| Child Protective Services (CPS) | | | | | |
| Referrals (children) Accepted for Investigation | 42,800 ¹ | 100% | \$618 ² | 2008 | \$618 |
| Police Involvement | 5,093 ³ | 11.9% | \$670 ⁴ | 2009 | \$80 |
| Juvenile Court Dependency Case Involvement | 3,883 ⁵ | 9.1% | \$354 ⁶ | 2009 | \$322 |
| Child Welfare Services | | | | | |
| Percentage of protective custody placements that are CPS cases | 75.27% ⁷ | | | | |
| Protective Custody (foster care) | 7,500 ¹ | 13.2% | \$24,568.50 ⁸ | 2008 | \$3,243 |
| In-Home Services (not out-of-home placement)* | 0 | 2008 | 0 | 2008 | \$0 |
| Adoption | 733 ⁹ | 1.7% | \$78,656.60 ¹⁰ | 2008 | \$1,337 |
| Juvenile Court Termination Case Involvement | 1,934 ⁵ | 4.5% | \$2,640 ⁶ | 2009 | \$119 |
| TOTAL: Expected present value cost of an accepted CPS case (in 2010 dollars) | | | | | \$5,719 |
| Addendum: Expected present value cost of an out-of-home placement, conditional on an out-of-home placement | | | | | \$32,936.64 |

Sources for Table D3.2c:

¹ Washington State DSHS Children's Administration, 2008 Performance Report, available at: <http://www.dshs.wa.gov/ca/pubs/2008perfrm.asp>

² Washington State DSHS Research and Data Analysis Client Data for FY2008. Total expenditures for "Child Protective Services case management", divided by total accepted referrals.

³ Percentage for Washington state from Administration on Children, Youth and Families (2008) Child Maltreatment 2008, Table 2-2, applied to total accepted referrals.

⁴ Marginal operating cost of an arrest for a misdemeanor from Institute crime model.

⁵ Washington State Office of the Administrator of the Courts, 2009, Juvenile dependency filings. Report available at <http://www.courts.wa.gov/caseload/content/archive/superior/Annual/atbls09.pdf>.

⁶ Based on average number of hearings per case (Miller, 2004) multiplied by WSIPP analysis of average cost per hearing (based on projected length in hours, and the hourly wages for the estimated number of people involved in each hearing).

⁷ Based on Institute analysis of DSHS Children's Administration data.

⁸ Calculated based on RDA data from FY2008: Total cost for out of home care, divided by total number of children in paid or relative care, multiplied by average length of stay out-of-home. Length of stay computed from entry and exit data for Washington in 2009 AFCARS report, available at http://www.acf.hhs.gov/programs/cb/stats_research/afcars/statistics/entryexit2009.htm.

⁹ Institute estimate of new adoption cases each year, from FY2008 DSHS Children's Administration data.

¹⁰ Institute calculation of total adoption support per case, estimated from FY2008 Children's Administration data.

* The cost of in-home services is not yet computed. The new data system for Washington State's child welfare system (FamLink) will have the ability to report that information in the future.

Expected value victim costs are derived from calculations by Miller, Fisher, and Cohen, 2001; their comprehensive analysis of the future impacts of victimization by child abuse and neglect takes into account medical, mental health, and quality of life costs, as described in Exhibit D3.2d below. We enter the summary taxpayer and victim costs on the input screen in Exhibit D3.1c. These estimated totals are life cycle expected value costs; we use the “decay” parameter for victim costs above to “spread out” those costs over a child’s life.

Exhibit D3.2d
Medical, Mental Health, and Quality of Life Costs per Victim of Child Abuse and Neglect
1993 Dollars

| | Medical and Mental Health Costs⁽¹⁾ | Quality of Life Costs⁽¹⁾ | Number of Victims⁽³⁾ |
|---|--|--|--|
| | (1) | (2) | (3) |
| Type of Child Abuse and Neglect | | | |
| Sexual abuse | \$6,327 ⁽²⁾ | \$94,506 ⁽²⁾ | 114,000 |
| Physical abuse | \$3,472 ⁽²⁾ | \$58,645 ⁽²⁾ | 308,000 |
| Mental abuse | \$2,683 ⁽²⁾ | \$21,099 ⁽²⁾ | 301,000 |
| Serious physical neglect | \$911 ⁽²⁾ | \$7,903 ⁽²⁾ | 1,236,000 |
| Total | \$1,901 ⁽⁴⁾ | \$22,948 ⁽⁴⁾ | 1,959,000 |
| Distribution of Costs by Payer | | | |
| Percentage incurred by taxpayer | 50% ⁽⁵⁾ | 0% ⁽⁵⁾ | |
| Percentage incurred by victim | 50% ⁽⁵⁾ | 100% ⁽⁵⁾ | |
| Amount paid by taxpayer | \$951 | \$0 | |
| Amount paid by victim | \$951 | \$22,948 | |
| Sources | | | |
| 1. The source of the cost elements in this table is Miller, T. R., Fisher, D. A., & Cohen, M. A. (2001). Costs of juvenile violence: Policy implications. <i>Pediatrics</i> , 107(1). DOI: 10.1542/peds.107.1.e3 | | | |
| 2. <i>Ibid.</i> , Table 1. We’ve assumed 80 percent urban and 20 percent rural costs on the Miller et al. Table 1. | | | |
| 3. The source for the total U.S. number of victims: Miller, T. R., Cohen, M. A., & Wiersema, B. (1996). <i>Victim costs and consequences: A new look</i> . Research report, Table 1. Washington, DC: National Institute of Justice. | | | |
| 4. These totals are weighted average sums using the victim numbers in column (3). | | | |
| 5. Institute assumptions. | | | |

D4. Valuation of Outcomes That Affect Alcohol and Illicit Drug Disorders, and Regular Tobacco Use

The Institute's benefit-cost model contains procedures to estimate the monetary value of changes in the disordered use of alcohol and illicit drugs, as well as the monetary value of changes in regular tobacco smoking. Illicit drugs represent a broad category of substances; the current version of the Institute's model divides illicit drugs into (a) cannabis and (b) all of other illicit drugs.⁸⁹ Analysts sometimes abbreviate alcohol, tobacco, and other drugs with the acronym ATOD. This section of the Technical Appendix describes the Institute's current procedures to estimate the monetary benefits of program-induced changes in ATOD. For the Institute's benefit-cost model, an alcohol and illicit drug disorder reflects both abuse and dependency as defined by the Diagnostic and Statistical Manual (DSM) of the American Psychiatric Association. Regular smoking is defined as daily smoking.

In general, analysts have constructed two types of studies to estimate the costs of ATOD: "prevalence-based" studies and "incidence-based" studies.⁹⁰ Prevalence costing studies look backward and ask: How much does ATOD cost society today, given all current and past disordered use of ATOD among people alive in a state or country? Incidence costing studies look forward and ask: How much benefit could be obtained in the future if disordered use of ATOD can be reduced? Both approaches use some of the same information, but assemble it different ways. Incidence-based studies are more useful for estimating the expected future benefits and costs of policy choices.

The Institute's ATOD model uses an incidence-based approach. Therefore, it is not designed to provide an estimate of the total cost to society of current and past ATOD. Other studies have attempted to estimate these values.⁹¹ For example, Rosen et al. found the total cost of alcohol in California in 2005 to be \$38.5 billion in "economic" costs (\$1,081 per capita) and an additional \$48.8 billion in "quality of life" costs.⁹² Similarly, Wickizer (2007) estimated the cost of alcohol to Washington State in 2005 to be \$2.9 billion in economic costs (\$466 per capita) and that illicit drugs cost Washington an additional \$2.3 billion.⁹³ These prevalence-based total cost studies can be interesting, but they are not designed to evaluate future marginal benefits and marginal costs of specific public policy options.

The purpose of the Institute's model is to provide the Washington State legislature with advice on whether there are economically attractive evidence-based policies that, if implemented well, can achieve reductions in the harmful use of ATOD. To do this, the model monetizes the projected life-cycle costs and benefits of programs or policies that have been shown to achieve improvements—today and in the future—in disordered ATOD. If, for example, empirical evidence indicates that a prevention program can delay the age at which young people initiate the use of alcohol, then what long-run benefits, if any, can be expected from this outcome? If an intervention program for current regular smokers can achieve a 10 percent reduction in the rate of smoking, then what are the life-course monetary benefits? Once computed, the present value of these benefits can be stacked against program costs to determine the relative attractiveness of different approaches to achieve improvements in desired outcomes.

The current version of the ATOD model allows the computation of the following types of avoided costs, or benefits, when a program or policy reduces probability of a person's current and future prevalence of ATOD. Depending on each particular ATOD, the following cost categories are included in the Institute's model:

- Labor market earnings from ATOD morbidity or mortality, to the degree there is evidence that current earnings are reduced because of ATOD (morbidity), or lifetime earnings are lost because of premature death (mortality) caused by ATOD, and that the program reduces the prevalence of ATOD.
- Medical costs for hospitalization and pharmaceuticals from ATOD morbidity or mortality, to the degree that these costs are caused by ATOD, and that the program reduces the prevalence of ATOD.
- Crime costs to taxpayers and victims, to the degree that crime is estimated to be caused by ATOD, and that the program reduces the prevalence of ATOD.
- Traffic collision costs, to the degree that collisions are estimated to be caused by ATOD, and that the program reduces the prevalence of ATOD.
- Value of a statistical life (VSL) estimates, net of labor market gains, applied to the change in mortality estimated to be caused by ATOD, and that the program reduces the prevalence of ATOD.

⁸⁹ Caulkins, J. P., & Kleiman, M. A. R. (n.d.). *Drugs and crime*. Unpublished manuscript, Carnegie Mellon University, Pittsburgh, PA.

⁹⁰ Moller, L., & Matic, S. (Eds.). (2010). *Best practice in estimating the costs of alcohol: Recommendations for future studies*. Copenhagen, Denmark: WHO Regional Office for Europe.

⁹¹ See, Harwood, H., Fountain, D., & Livermore, G. (1998). *The economic costs of alcohol and drug abuse in the United States 1992* (NIH Publication No. 98-4327). Rockville, MD: National Institutes of Health. See also, Rice, D. P., Kelman, S., Miller, L. S., & Dunmeyer, S. (1990). *The economic costs of alcohol and drug abuse and mental illness, 1985* (DHHS Pub. No.90-1694). Washington, DC: Alcohol, Drug Abuse, and Mental Health Administration.

⁹² Rosen, S. M., Miller, T. R., & Simon, M. (2008). The cost of alcohol in California. *Alcoholism: Clinical and Experimental Research*, 32(11), 1925-1936. The California study uses a few incidence-based methods in addition to prevalence-based methods.

⁹³ Wickizer, T. M. (2007, June). *The economic costs of drug and alcohol abuse in Washington State, 2005*. Olympia: Washington State Department of Social and Health Services, Division of Alcohol and Substance Abuse.

D4.1 Input Screens for ATOD Parameters

The ATOD model is driven with a set of parameters describing various aspects of ATOD epidemiology and linked relationships with other outcomes. These input parameters are shown in the following four screen shots. In addition, there are several other input parameters used in the ATOD model that are general to the Institute's overall benefit-cost model, and these are discussed elsewhere in this Appendix. In the following sections, the sources for the parameters and the computational routines are described.

Exhibits D4.a through D4.d display the parameters for the analysis of disordered alcohol, tobacco, cannabis, and other illicit drug use.

Exhibit D4.a Alcohol Disorders

WSIPP Benefit-Cost Model: Version 1.2

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

General | Economic | Crime | Education | Child Welfare | **Substance Use** | Health Care | Mental Health | Public Asst | Housing | Teen Birth | Outcomes & Links

Substance Use (ATOD) [X]

Back to Main Model

Alcohol | Tobacco | Cannabis | Other Illicit Drugs

DSM Alcohol Use Disorders--Epidemiology

Proportion of general population with lifetime alcohol use disorder:

Age of Onset of DSM Alcohol Disorders: the three parameters for a LogLogistic probability density distribution.

Remission Rate: parameters for a Weibull distribution. (We use the inverse to describe persistence of the disorder.)

shift alpha beta

Proportion of general population that consumes alcohol:

Standard deviation (yrs) in the age of initiation:

Annual Alcohol Attributed Deaths

| Age group | Number of years in age group | Alcohol attributed deaths: Chronic | Alcohol attributed deaths: Acute | Proportion of acute deaths attributable to DSM Alcohol disorder | State deaths (all) | State population in age group |
|-----------|------------------------------|------------------------------------|----------------------------------|---|--------------------|-------------------------------|
| 1-19 | 19 | 2 | 57 | 0.5 | 891.8 | 1699651 |
| 20-34 | 15 | 11 | 220 | 0.5 | 1020.8 | 1263739 |
| 35-49 | 15 | 177 | 259 | 0.5 | 3120.3 | 1433694 |
| 50-64 | 15 | 291 | 149 | 0.5 | 6374.5 | 1022490 |
| 65+ | 36 | 301 | 218 | 0.5 | 33858.2 | 688250 |

The Year(s) these data represent:

DSM Alcohol Use Disorders: Monetary Consequences

Labor Market parameters

Gain in labor market earnings for never alcoholics vs current alcoholics, lognormal probability density distribution parameters

| Mean | Std dev |
|-------------------------------------|------------------------------------|
| <input type="text" value="0.1389"/> | <input type="text" value="0.062"/> |

Gain in labor market earnings for former alcoholics vs current alcoholics, lognormal probability density distribution parameters

| Mean | Std dev |
|-------------------------------------|------------------------------------|
| <input type="text" value="0.1389"/> | <input type="text" value="0.062"/> |

Hospital-related Parameters

| | | | |
|------------------------------------|---|--|-----------------|
| <input type="text" value="16505"/> | Annual number of DO FTE hospital events | <input type="text" value="2007"/> | Year of data |
| <input type="text" value="24515"/> | Avg charge per DO FTE event | <input type="text" value="2007"/> | Year of dollars |
| <input type="text" value="4.88"/> | Number of days per DO FTE stay | FTE: full time equivalent disorder event | |

Emergency Department-related Parameters

| | | | |
|------------------------------------|--|-----------------------------------|-----------------|
| <input type="text" value="0.079"/> | Proportion of admissions attributable to alcohol | | |
| <input type="text" value="569"/> | ER charge per admission, dollars | <input type="text" value="2008"/> | Year of dollars |

Treatment Parameters

| | | | |
|------------------------------------|--------------------------------------|-----------------------------------|--------------------------------|
| <input type="text" value="15777"/> | Annual number treated | <input type="text" value="2010"/> | Year of data |
| <input type="text" value="1551"/> | Cost per treatment episode | <input type="text" value="2005"/> | Year of data |
| <input type="text" value="0"/> | Percent cost paid by self | <input type="text" value="1"/> | Percent cost paid by taxpayers |
| <input type="text" value="0"/> | Percent cost paid by private insurer | | |

Traffic Crash-related Parameters

| | | | |
|------------------------------------|---------------------------------------|-----------------------------------|------------------------------|
| <input type="text" value="15381"/> | Annual number alcohol-related crashes | <input type="text" value="2009"/> | Year of data |
| <input type="text" value="1891"/> | Avg property cost per crash | <input type="text" value="2000"/> | Year of dollars |
| <input type="text" value="0.35"/> | Percent cost paid by self | <input type="text" value="0.65"/> | Percent cost paid by insurer |

Exhibit D4.b Regular Smoking

WSIPP Benefit-Cost Model: Version 1.2

Enter Sector Inputs
Enter Program Inputs
Run Models & View Reports

General

Economic

Crime

Education

Child Welfare

Substance Use

Health Care

Mental Health

Public Asst

Housing

Teen Birth

Outcomes & Links

Substance Use (ATOD)

[Back to Main Model](#)

Alcohol
Tobacco
Cannabis
Other Illicit Drugs

Regular Tobacco Smoking--Epidemiology

Proportion of general population with lifetime regular tobacco smoking. 0.393

Proportion of general population that smokes tobacco. 0.278

Age of Onset of regular tobacco smoking: the three parameters for a LogLogistic probability density distribution.

12.647 Standard deviation (yrs) in the age of initiation. 3.3

6.8346 3.3

Remission Rate: shift 0.5

parameters for a Beta distribution. (We use the inverse to describe persistence of the disorder.)

alpha 0.96399

beta 2.0358

lower bound 0

upper bound 115.25

Regular Tobacco Smoking: Monetary Consequences

Labor Market parameters

Gain in labor market earnings for never smokers vs current regular smokers, lognormal probability density distribution

| | | |
|--|--------|---------|
| | Mean | Std dev |
| | 0.1085 | 0.0904 |

Gain in labor market earnings for former smokers vs current regular smokers, lognormal probability density distribution

| | | |
|--|---|--------|
| | 0 | 0.0001 |
|--|---|--------|

Hospital-related Parameters

| | | | |
|-------|---|------|--|
| 22914 | Annual number of DO FTE hospital events | 2007 | Year of data |
| 35671 | Avg charge per DO FTE event | 2007 | Year of dollars |
| 4.36 | Number of days per DO FTE stay | | FTE: full time equivalent disorder event |

Emergency Department-related Parameters

| | | | |
|-------|--|------|-----------------|
| 0.049 | Proportion of admissions attributable to tobacco | | |
| 569 | ER charge per admission, dollars | 2008 | Year of dollars |

Treatment Parameters

| | | | |
|---|--------------------------------------|------|--------------------------------|
| 0 | Annual number treated | 2010 | Year of data |
| 0 | Cost per treatment episode | 2005 | Year of data |
| 0 | Percent cost paid by self | 1 | Percent cost paid by taxpayers |
| 0 | Percent cost paid by private insurer | | |

Annual Tobacco Smoking Attributed Deaths

| Age group | Number of years in age group | Smoking attributed deaths | State deaths (all) | State population in age group |
|-----------|------------------------------|---------------------------|--------------------|-------------------------------|
| 1-34 | 34 | 0 | 1,991 | 3113578 |
| 35-44 | 10 | 121.75 | 1,330 | 914832 |
| 45-54 | 10 | 537.81 | 3,524 | 983194 |
| 55-64 | 10 | 1257.23 | 5,864 | 770691 |
| 65-74 | 10 | 1583.44 | 7,571 | 417524 |
| 75-84 | 10 | 2263.98 | 12,368 | 256598 |
| 85-100 | 16 | 1456.34 | 15,902 | 109656 |

The Year(s) these data represent: 2008

68

Exhibit D4.c Cannabis Disorders

WSIPP Benefit-Cost Model: Version 1.2 X

Enter Sector Inputs
Enter Program Inputs
Run Models & View Reports

General

Economic

Crime

Education

Child Welfare

Substance Use

Health Care

Mental Health

Public Asst

Housing

Teen Birth

Outcomes & Links

Substance Use (ATOD) X

Back to Main Model

Alcohol
Tobacco
Cannabis
Other Illicit Drugs

Disordered Cannabis Use--Epidemiology

Proportion of general population with lifetime cannabis disorder. 0.085

Age of Onset of Cannabis Disorders; the three parameters for an Extreme Value probability density distribution. 18.0348

Age of Onset of Cannabis Disorders; the three parameters for an Extreme Value probability density distribution. 3.6638

Remission Rate: parameters for a Lognormal distribution. (We use the inverse to describe persistence of the disorder.)

| | | |
|--|------|--------|
| | mean | 1.7917 |
| | sd | 1.149 |

Proportion of general population that consumes cannabis. 0.114

Standard deviation (yrs) in the age of initiation. 3.6

Disordered Cannabis Use: Monetary Consequences

Labor Market parameters

| | Mean | Std dev |
|--|--------|---------|
| Gain in labor market earnings for never used cannabis vs current disordered users, lognormal probability density distribution parameters | 0.0427 | 0.01 |
| Gain in labor market earnings for former users vs current disordered users, lognormal probability density distribution parameters | 0.0427 | 0.01 |

Treatment Parameters

| | | | |
|------|--------------------------------------|------|--------------------------------|
| 8524 | Annual number treated | 2010 | Year of data |
| 1551 | Cost per treatment episode | 2005 | Year of data |
| 0 | Percent cost paid by self | 1 | Percent cost paid by taxpayers |
| 0 | Percent cost paid by private insurer | | |

Exhibit D4.d Other Illicit Drug Disorders

WSIPP Benefit-Cost Model: Version 1.2 X

Enter Sector Inputs
Enter Program Inputs
Run Models & View Reports

General

Economic

Crime

Education

Child Welfare

Substance Use

Health Care

Mental Health

Public Asst

Housing

Teen Birth

Outcomes & Links

Substance Use (ATOD)
X

Back to Main Model

Alcohol
Tobacco
Cannabis
Other Illicit Drugs

Disordered Illicit Drug Use--Epidemiology

Proportion of general population with lifetime illicit drug disorder. 0.055

Age of Onset of Other Illicit Drug Disorders; the three parameters for an Extreme Value probability density distribution. 18.0348

Age of Onset of Other Illicit Drug Disorders; the three parameters for an Extreme Value probability density distribution. 3.6638

Remission Rate: parameters for a Lognormal distribution. (We use the inverse to describe persistence of the disorder.)

| | | |
|------|--------|--|
| mean | 1.4741 | |
| sd | 1.0985 | |

Proportion of general population that consumes illicit drugs. 0.084

Standard deviation (yrs) in the age of initiation. 4.18

Annual Illicit Drug Disorder Attributed Deaths

| | Lower age | Upper age | Drug attributed deaths | State deaths (all) | State population in age group |
|--|-----------|-----------|------------------------|--------------------|-------------------------------|
| | 1 | 14 | 7.5 | 615.6 | 1251485 |
| | 15 | 19 | 17.0 | 243.2 | 432244.2 |
| | 20 | 24 | 22.0 | 355.6 | 434752 |
| | 25 | 34 | 29.5 | 713.4 | 869927.6 |
| | 35 | 44 | 39.5 | 1453 | 939210.8 |

The Year(s) these data represent: 2003-07

Disordered Illicit Drug Use: Monetary Consequences

Labor Market parameters

| | | | |
|--|--------|---------|--|
| | Mean | Std dev | |
| Gain in labor market earnings for never abusers vs current abusers, lognormal probability density distribution parameters | 0.0427 | 0.01 | |
| Gain in labor market earnings for former abusers vs current abusers, lognormal probability density distribution parameters | 0.0427 | 0.01 | |

Hospital-related Parameters

| | | | |
|-------|---|------|--|
| 9730 | Annual number of DO FTE hospital events | 2007 | Year of data |
| 23566 | Avg charge per DO FTE event | 2007 | Year of dollars |
| 5.16 | Number of days per DO FTE stay | | FTE: full time equivalent disorder event |

Emergency Department-related Parameters

| | | | |
|--------|--|------|-----------------|
| 0.0084 | Proportion of admissions attributable to illicit drugs | | |
| 569 | ER charge per admission, dollars | 2008 | Year of dollars |

Treatment Parameters

| | | | |
|-------|--------------------------------------|------|--------------------------------|
| 13138 | Annual number treated | 2010 | Year of data |
| 1551 | Cost per treatment episode | 2005 | Year of dollars |
| 0 | Percent cost paid by self | 1 | Percent cost paid by taxpayers |
| 0 | Percent cost paid by private insurer | | |

D4.2 ATOD Epidemiological Parameters: Current Prevalence for Prevention and Intervention Programs

The Institute's ATOD model begins by analyzing the epidemiology of each ATOD disorder to produce estimates of the current 12-month prevalence of disordered alcohol use, disordered illicit drug use, and regular tobacco smoking. An estimate of the current prevalence of an ATOD disorder is central to the benefit-cost model because it becomes the "base rate" of an ATOD disorder to which program or policy effect sizes are applied to calculate the change in the number of avoided ATOD "units" caused by the program, over the lifetime following treatment.

Four parameters enter the model to enable an estimate of the current prevalence of ATOD, from age 1 to age 100.

- **Lifetime prevalence:** the percentage of the population that has a specific lifetime ATOD disorder.
- **Age of onset:** the age of onset of the specific ATOD disorder.
- **Persistence:** the persistence of the specific ATOD disorder, given onset.
- **Death (Survival):** the probability of death by age, after the age of treatment by a program.

The parameters that enter the model appear on each screen shot; Exhibit D4.a also displays the current parameters in the Institute's model for the first three epidemiological factors, along with sources and notes. The death probability information is described elsewhere in this Appendix.

For each ATOD disorder, the current prevalence of ATOD is estimated with this equation.

$$(D.4.1) \quad CP_y = \left(\sum_{0=1}^y O_0 \times P_{(y-0+1)} \right) \times LTP \times S_y$$

The current disorder prevalence probability at any year in a person's life, CP_y , is computed with information on the age-of-onset probability, O_y , from prior ages to the current age of the person, times the persistence probability, P , of remaining in the DSM condition at each onset age until the person is the current age, times the lifetime probability of ever having the DSM disorder (or regular tobacco use), LTP , times the probability of survival at each age, S_y , following treatment by a program.

For each ATOD disorder, the exogenous age-of-onset probability distribution for ages 1 to 100, O_y , is a density distribution and is estimated with information from the sources shown in Exhibit D4.e. The parameters in Exhibit D4.e are the same as those entered by the user on the screen shots in Exhibits D4.a through D4.d.

$$1 = \sum_{y=1}^{100} O_y$$

Also, for each ATOD disorder, the exogenous persistence distribution for ages after onset, P , is computed from the sources shown in Exhibit D4.e. The persistence distribution describes the probability, on average, of being in the DSM disorder condition each year following onset.

The probability of survival at any given age, S_y , is computed from a national life table on survival, LTS , in the general population. The inputs for the survival table are described in another section of this Technical Appendix. To compute the current prevalence of a disorder over the entire life course, S_y is normalized to age 1.

$$S_y = \frac{LTS_y}{LTS_1}$$

Since the probability of survival depends on the number still living at the treatment age, ta_{ge} , the S_y is normalized to the age of the person being treated in the program being analyzed, since it is assumed that all treatment programs will be for those currently alive at time of treatment.

$$S_y = \frac{LTS_y}{LTS_{ta_{ge}}}$$

Equation D.4.1 describes the calculation of current prevalence for general (prevention) populations. For programs treating indicated populations, CP_y in equation D.4.2 describes the prevalence in all years following treatment.

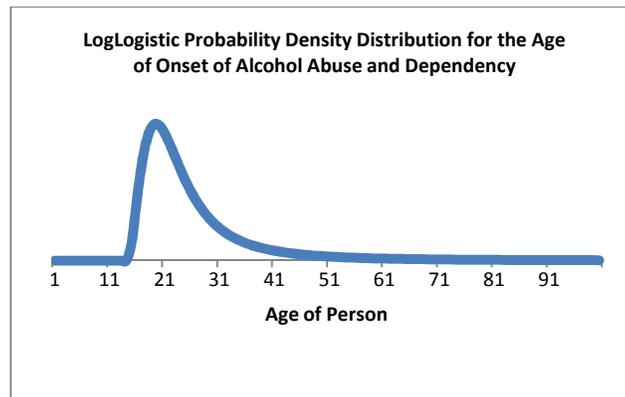
$$(D.4.2) \quad CP_y = \frac{\sum_{0=1}^{tage} 0_0 \times P_{(y-0+1)}}{\sum_{0=1}^{tage} 0_0} \times S_y \times SF$$

The additional term in equation D.4.2 is the reduced chance of survival for someone with an ATOD disorder. We compute an estimate for this as a single parameter with the following equation.

$$SF = \frac{\sum_{a=1}^A \left(Pop_a \times CP_a \times \frac{(PopD_a - AtodD_a)}{\frac{PopD_a}{Pop_a}} \right)}{\sum_{a=1}^A (Pop_a \times CP_a)}$$

In this equation, Pop_a is the total population in a state in each age group, CP_a is the average current ATOD prevalence in each age group, $PopD_a$ is the total number of deaths in a state in each age group, and $AtodD_a$ is the deaths attributable to ATOD in each age group.

Example. We provide an illustrative example of computing CP_y in equation D.4.1 for alcohol disorders. Using the results from Hasin et al., we computed a probability density distribution for the age of onset of DSM alcohol disorders.⁹⁴ The Hasin study summarizes information from the National Epidemiologic Survey on Alcohol and Related Conditions, a nationally representative sample. We used *@Risk* software to estimate alternative distributions that fit the onset information reported in the Hasin study. We then selected the type of distribution with the best fit where the criterion was the lowest root-mean squared error. For our analysis of the results reported in the Hasin study, we computed a loglogistic density distribution; the estimated parameters are reported in Exhibit D4.e. The chart below plots the estimated distribution, where the sum of annual probabilities equals 1.0



Next, estimates of the persistence of the alcohol disorder, given onset, were computed for alcohol from the study by Lopez-Quintero, et al.⁹⁵ The Lopez-Quintero study also used information from the National Epidemiologic Survey on Alcohol and Related Conditions. Again, we used *@Risk* software to model the best fitting cumulative remission curve, and then inverted the result to estimate a persistence curve. A Weibull distribution was the best-fitting curve for this disorder. The resulting estimates measure the probability of remaining in a DSM alcohol disorder in the years following onset. The estimated Weibull parameters are shown in Exhibit D4.e and the chart below plots the results.

⁹⁴ Hasin, D. S., Stinson, F. S., Ogburn, E., & Grant, B. F. (2007). Prevalence, correlates, disability and comorbidity of DSM-IV alcohol abuse and dependence in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, 64(7), 830-842.

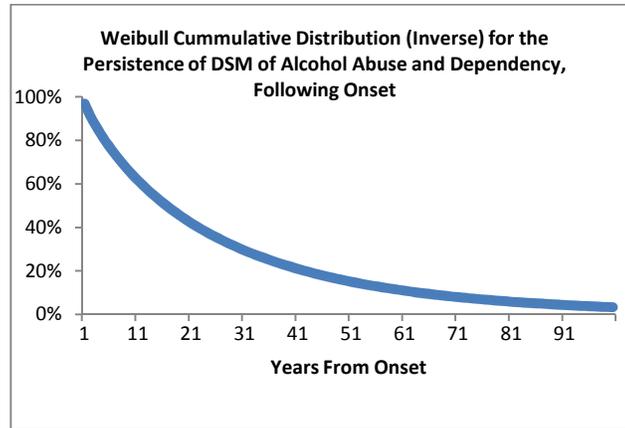
⁹⁵ Lopez-Quintero, C., Hasin, D. S., de los Cobos, J. P., Pines, A., Wang, S., Grant, B. F., & Blanco, C. (2011). Probability and predictors of remission from lifetime nicotine, alcohol, cannabis, or cocaine dependence: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Addiction*, 106(3), 657-669.

Exhibit D4.e
Input Parameters for the Epidemiology of Alcohol Disorders, Illicit Drug Disorders, and Regular Smoking⁽¹⁾

| | DSM Alcohol Disorder | DSM Illicit Drug Disorder (Cannabis) | DSM Illicit Drug Disorder (Non Cannabis) | Regular Tobacco Smoking |
|---|-----------------------------|--------------------------------------|--|------------------------------|
| | (a) | (b) | (c) | (d) |
| Percentage of population with lifetime DSM disorder, or regular smoking | 24.2% ⁽²⁾ | 8.5% ⁽⁸⁾ | 5.5% ⁽⁸⁾ | 39.3% ⁽¹¹⁾ |
| Age of onset | | | | |
| Type of distribution | Log-logistic ⁽³⁾ | Extreme value ⁽⁹⁾ | Extreme value ⁽⁹⁾ | Log-logistic ⁽¹²⁾ |
| Parameter 1 | 14.5776 | 18.0348 | 18.0348 | 4.5788 |
| Parameter 2 | 8.0661 | 3.6638 | 3.6638 | 12.647 |
| Parameter 3 | 2.05 | n/a | n/a | 6.8346 |
| Parameter 4 | n/a | n/a | n/a | n/a |
| Remission of DSM disorder, given onset | | | | |
| Type of distribution | Weibull ⁽⁴⁾ | Lognormal ⁽⁴⁾ | Lognormal ⁽⁴⁾ | Beta-general ⁽⁴⁾ |
| Parameter 1 | .5 | 1.7917 | 1.4741 | .5 |
| Parameter 2 | .86728 | 1.149 | 1.0985 | .96399 |
| Parameter 3 | 24.119 | n/a | n/a | 2.0358 |
| Parameter 4 | n/a | n/a | n/a | 0 |
| Parameter 5 | n/a | n/a | n/a | 115.25 |
| Percentage of general population consuming substance | 67.2% ⁽⁵⁾ | 11.4% ⁽⁵⁾ | 8.4% ⁽⁵⁾ | 27.8% ⁽⁵⁾ |
| Age of initiation parameters | | | | |
| Standard deviation in age of initiation (years) | 3.32 ⁽⁶⁾ | 3.60 ⁽⁶⁾ | 4.18 ⁽⁶⁾ | 3.30 ⁽⁶⁾ |
| Effect Size: current DSM prevalence per year of delay | .020 ⁽⁷⁾ | .050 ⁽¹⁰⁾ | .024 ⁽¹⁰⁾ | .025 ⁽¹³⁾ |
| Standard Error | .019 ⁽⁷⁾ | .011 ⁽¹⁰⁾ | .009 ⁽¹⁰⁾ | .028 ⁽¹³⁾ |

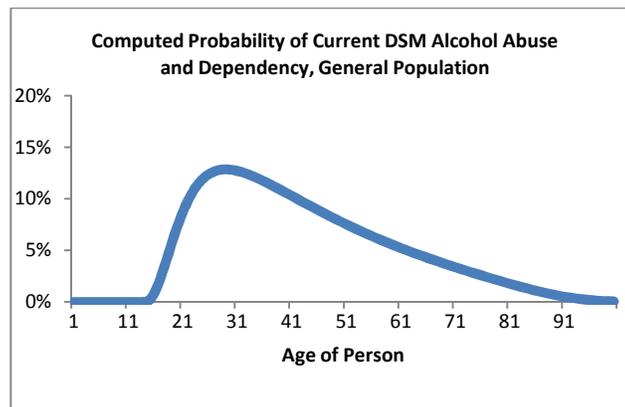
Notes and sources

- For benefit-cost modeling, except where noted, alcohol and drug disorders include both DSM categories of abuse and dependence. Tobacco smoking is measured as regular daily smoking. All outcomes are estimated as dichotomous conditions.
- Vergés, A., Littlefield, A. K., & Sher, K. J. (2001, January 10). Did lifetime rates of alcohol use disorders increase by 67% in ten years? A comparison of NLAES and NESARC. *Journal of Abnormal Psychology*. Advance online publication. This study compares results from the NLAES and NESARC epidemiological surveys. We elected to average the two results for the two national surveys reported in the Vergés study (.1817 and .3028). When the averaged lifetime value is entered into our model, the resulting current prevalence estimate from our model (.077) is nearly identical to the average of the current prevalence estimates, reported by Vergés, from the two national surveys (.079, the average of .0740 and .0846).
- Hasin, D. S., Stinson, F. S., Ogburn, E., & Grant, B. F. (2007). Prevalence, correlates, disability and comorbidity of DSM-IV alcohol abuse and dependence in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, 64(7), 830-842. From the Figure reported in the paper, we computed a loglogistic probability density distribution for the age of onset of a DSM alcohol disorder, conditional on having a disorder. @Risk software was used to estimate alternative distributions; the distribution with the best fit (criterion: lowest root-mean squared error) was chosen.
- Lopez-Quintero, C., Hasin, D. S., de los Cobos, J. P., Pines, A., Wang, S., Grant, B. F., & Blanco, C. (2011). Probability and predictors of remission from lifetime nicotine, alcohol, cannabis, or cocaine dependence: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Addiction*, 106(3): 657-669. For alcohol and illicit drug disorders and nicotine we fitted cumulative probability distributions to the remission information reported in the study, and then inverted to estimate persistence curves. @Risk software was used to estimate alternative distributions; for each disorder, the distribution with the best fit (criterion: lowest root-mean squared error) was chosen. For alcohol and tobacco, the first parameter shown is a shift parameter. For illicit drug disorders, the non-cannabis estimate is for cocaine, the only non-cannabis illicit drug analyzed in the Lopez-Quintero paper.
- Analysis of 2009 National Survey on Drug Use and Health. For alcohol, we used the ALCYR variable (used within the past year). We used the MRJYR variable for cannabis (used in past year), the IEMYR variable for illicit drugs other than cannabis (used in past year), and the CIGYR variable (used in past year) for cigarettes.
- Analysis of 2009 National Survey on Drug Use and Health. For alcohol, we used the IRALCAGE variable (age of initiation, filtered for initiation ages 10 to 30—for analysis of prevention programs, age of initiation beyond 30 is not relevant). We used the IRMJAGE variable (age of cannabis initiation) and IEMAGE (age of initiation of illicit drug use other than cannabis), both filtered for initiation ages 10 to 30—for analysis of prevention programs, age of initiation beyond 25 is not relevant. For cigarettes, we used the IRCIGAGE variable (age of initiation, filtered for initiation ages 7 to 25—for analysis of prevention programs, age of initiation beyond 25 is not relevant).
- These parameters were computed from an analysis of the research literature examining the probability of the current prevalence of adult DSM alcohol disorder as a function of age of initiation of alcohol consumption. In the analysis, we contributed our own study using the 2009 NSDUH dataset. The units shown are effect sizes on adult DSM alcohol disorders (and standard errors) per year of delay in initiation. The mean effect size was reduced by half to be consistent with the Institute's adjustments for unobserved selection bias.
- Compton, W. M., Thomas, Y. F., Stinson, F. S., Grant, B. F. (2007). Prevalence, correlates, disability and comorbidity of DSM-IV drug abuse and dependence in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, 64(5), 566-576. Cannabis disorder prevalence reported in eTable 1. The Compton paper did not report a separate estimate for lifetime prevalence for non-cannabis illicit drugs. We estimated this by applying the data from the 2009 NSDUH, multiplying the current non-cannabis illicit drug prevalence (ABODIEM) by the ratio of lifetime cannabis illicit drug prevalence from the Compton paper to current cannabis prevalence (ABODMRJ) from the NSDUH.
- Ibid*. From the Figure reported in the Compton paper, we computed an extreme value probability density distribution for the age of onset of a DSM drug disorder, conditional on having a disorder. @Risk software was used to estimate alternative distributions; the extreme value distribution fit the Compton data well, especially for early ages. The Compton study only reported distributions for all drugs, not separate curves for cannabis and non-cannabis illicit drugs. Hence, we use the same density distribution for both cannabis and other illicit drugs; future research can refine this.
- These parameters were computed from a multivariate logistic regression analysis of 2009 National Survey on Drug Use and Health data where the probability of the current prevalence of a DSM cannabis use disorder (ABODMRJ) was related to the age of onset of cannabis use (IRMJAGE). Covariates included current age, gender, income, and race. The units shown are effect sizes (and standard errors) per year of delay in initiation. We conducted a similar analysis for DSM non-cannabis illicit drug use disorder. Variables used were ABODIEM, IEMAGE, and current age, gender, income, and race covariates. Mean effect sizes were reduced by half to be consistent with the Institute's adjustments for unobserved selection bias.
- Analysis of 2009 National Survey on Drug Use and Health. We used the CIGDLYMO variable (ever smoked cig every day for 30 days) and filtered for ages 26 to 49 to match a post initiation cohort and a post-surgeon general's cohort.
- Analysis of 2009 National Survey on Drug Use and Health. We used the IRCDUAGE variable (imputation-revised daily cig age of first use). We computed a log-logistic probability density distribution for the age of onset of regular cigarette use. @Risk software was used to estimate alternative distributions; the distribution with the best fit (criterion: lowest root-mean squared error) was chosen.
- These parameters were computed from an analysis of the research literature examining the probability of the current prevalence of adult regular smoking as a function of age of initiation of smoking. In the analysis, we contributed our own study using the 2009 NSDUH dataset. The units shown are effect sizes on adult regular smoking (and standard errors) per year of delay in initiation. The mean effect size was reduced by half to be consistent with the Institute's adjustments for unobserved selection bias.



The persistence curve, after multiplying by the survival factor, by year, from the 2006 United States life table published by the federal Center for Disease Control, supplies the base rates for intervention programs.

For prevention programs, after applying the estimate of lifetime prevalence of an alcohol disorder, 24.2 percent with sources shown in Exhibit D4.e, and after adjusting for survival from the 2006 United States life table published by the federal Center for Disease Control (and assuming for this example a treatment age of one), the expected current 12-month prevalence of an alcohol disorder during the lifetime of a general population of one-year-olds is computed with equation D.4.1 and is plotted on the following chart.



The same procedures just described for alcohol disorders are used for disordered illicit drug use (non-cannabis), DSM Cannabis use, and regular tobacco smoking, substituting the relevant parameters for the best-fitting distributions as shown in Exhibit D4.e. As noted, the estimates of the current prevalence of an ATOD is central to the benefit-cost model because it becomes the “base rate” of an ATOD disorder to which program or policy effect sizes are applied to determine the change in the number ATOD “units” caused by the program, over the lifetime following treatment. The general prevalence, shown above, is used for programs targeted at the general population, while the persistence curve (after adjustment for survival probabilities), also shown above, is used as the base rate for programs that treat people with a current ATOD condition.

D4.3 ATOD Attributable Deaths

The Institute’s model computes mortality-related lost earnings, lost household production, and the value of a statistical life. These mortality estimates require estimates of the probability of dying from ATOD. The model inputs for these calculations, for each ATOD disorder, are shown in Exhibits D4.a for alcohol, D4.b for smoking, and D4.d for illicit drugs other than cannabis.

Alcohol. For alcohol-attributable deaths, the data source is the United States Department of Health and Human Services, Center for Disease Control (CDC). CDC has estimated, for each state, the number of deaths attributable to alcohol causes.

The estimates from CDC are available on-line via a software application called *Alcohol-Related Disease Impact (ARDI)*.⁹⁶ According to CDC:

ARDI either calculates or uses pre-determined estimates of Alcohol-Attributable Fractions (AAFs)—that is, the proportion of deaths from various causes that are due to alcohol. These AAFs are then multiplied by the number of deaths caused by a specific condition (e.g., liver cancer) to obtain the number of alcohol-attributable deaths.

A Scientific Work Group, comprised of experts on alcohol and health, was convened to guide development of the ARDI software. The Work Group's tasks included:

- * Selecting alcohol-related conditions to be included in the application*
- * Selecting relative risk estimates for the calculation of alcohol-attributable fractions for specific conditions*
- * Determining prevalence cutpoints for different levels of alcohol use*

The most recent CDC/ARDI estimates for Washington State are the average annual number of alcohol attributable deaths, by age group shown of Exhibit D4.a, for the years 2001-05. ARDI estimates deaths related entirely or partially due to particular causes of death. For the deaths partially caused by alcohol, we obtained only the deaths associated with the ARDI “medium and high” alcohol consumption levels, since problem drinking is the focus of our benefit-cost analysis. ARDI also reports deaths due to chronic conditions (e.g. liver cirrhosis, fetal alcohol syndrome, etc.) and acute conditions (e.g. fall injuries, motor vehicle crashes, etc.). Since the Institute’s model focuses on DSM-level alcohol disorders, a portion of the deaths caused by acute conditions could be from alcohol-involved events of someone not with a DSM-level condition. Therefore, for acute deaths, the input screen provides for a single parameter, by age group, to split acute alcohol-related deaths into those where a DSM-alcohol disordered person was involved.

To compute alcohol induced death rates for these age groups, we obtained Washington State population data from the Washington State Office of Financial Management, the state agency charged with compiling official state demographic data. The population estimates are the average Washington population for 2001-05, the same years as the CDC/ARDI death estimates.

Tobacco Smoking. For smoking-attributable deaths, the data source is also the United States Department of Health and Human Services, Center for Disease Control. CDC has estimated, for each state, the number of deaths attributable to smoking. The estimates from CDC are available on-line via a software application called *Smoking-Attributable Mortality, Morbidity, and Economic Costs (SAMMEC)*.⁹⁷ SAMMEC reports smoking-attributable fractions of deaths for 19 diseases where cigarette smoking is a cause using sex-specific smoking prevalence and relative risk (RR) of death data for current and former smokers aged 35 and older.

Illicit Drugs. For illicit drug deaths, we used death data from the Washington State Vital Statistics dataset for the years 2003 to 2007. For these years, we counted the age of all deaths in Washington where ICD-10 death codes matched the drug attribution factors contained in Harwood et al.⁹⁸ We computed average annual drug attributable deaths in the age groups shown in Exhibit D4.d.

For each ATO, the death data are used to compute the probability of dying from ATOD in the general population, by age group.

$$AtodD_a = ((Chronic_a + Acute_a \times AcutePct) / Pop_a) / Years_a$$

The probability of dying from a particular ATOD disorder in each age group in the general population, $AtodD_a$, is computed by adding the deaths due to chronic ATOD use, $Chronic_a$, to the proportion of deaths due to acute ATOD use (e.g., motor vehicle crashes due to an alcohol impaired driver), $Acute_a$ times $AcutePct_a$, divided by the total population in the state in each age group, Pop_a . This quotient is divided by the number of years in the age group, $Years_a$, to produce an estimate of the average annual probability of dying from an ATOD disorder.

⁹⁶ Centers for Disease Control and Prevention website: <https://apps.nccd.cdc.gov/ardi/HomePage.aspx>

⁹⁷ Centers for Disease Control and Prevention website: <http://apps.nccd.cdc.gov/sammecl/>

⁹⁸ Office of National Drug Control Policy. (2004, December). *The economic costs of drug abuse in the United States, 1992-2002* (Publication No. 207303). Washington, DC: Executive Office of the President, Author, Table B-10.

D4.4 Linkages: ATOD and Other Outcomes

The Institute’s benefit-cost model monetizes improvements in ATOD outcomes, in part, with linkages between each ATOD and other outcomes to which a monetary value can be estimated. The parameters for these linkages are obtained by a meta-analytic review of relevant research literature. For example, we estimate the relationship between disordered alcohol use and labor market earnings by meta-analyzing the most credible studies that have addressed this topic. The meta-analytic process provides both an expected value effect given the weight of the evidence, and an estimate of the error of the estimated effect. Both of these two parameters are entered into the benefit-cost model and used when performing a Monte Carlo simulation. The linkages in the current Institute model are listed in Appendix E.

D4.5 Human Capital Outcomes Affecting Labor Market Earnings via ATOD-Caused Morbidity and Mortality

The Institute model computes lost labor market earnings, as a result of ATOD morbidity and mortality, when there is evidence that the linkage is causal. The procedures begin by estimating the labor market earnings of an average person with a current ATOD disorder. As described in Appendix D.1, the Institute’s model uses national earnings data from the U.S. Census Bureau’s Current Population Survey. The CPS data used in this analysis represent average earnings of all people, both workers and non-workers at each age.

For each person at each age, total CPS earnings can be viewed as a weighted sum of people who have never had an ATOD disorder, plus those that are currently disordered, plus those that were formerly disordered. From the CPS data on total earnings for all people, the earnings of individuals with a current ATOD condition, at each age, y , is computed with this equation:

$$EarnC_y = \frac{EarnAll_y \times (1 + EarnEscAll)^{y-tage} \times EarnBenAll \times (1 + EarnBenEscAll)^{y-tage} \times (IPD_{base}/IPD_{cps})}{\left((1 + EarnGN) \times \left(1 - \left(CP_y + \left(\sum_{o=1}^y (O_o \times LTP) - CP_y \right) \right) \right) + (1 + EarnGF) \times \left(\sum_{o=1}^y (O_o \times LTP) - CP_y \right) + CP_y \right)}$$

The numerator in the above equation includes the CPS earnings data for all people, $EarnAll$, with adjustments for real earnings growth, $EarnEscAll$, earnings-related benefits, $EarnBenAll$, growth rates in earnings benefits, $EarnBenEscAll$, and an adjustment to denominate the year of the CPS earnings data, IPD_{cps} , with the year chosen for the overall analysis, IPD_{base} . These variables are described in Appendix D.1.

The denominator uses the epidemiological variables described above: age of onset probabilities, O_y , lifetime prevalence rates, LTP_y , and current 12-month prevalence rates, CP_y , at each age.

The denominator also includes two variables on the earnings gain of never-disordered people compared to currently disordered people, $EarnGN$, and the earnings gain of formerly disordered people compared to currently disordered people, $EarnGF$. These two central relationships measure the effect of ATOD on labor market success (as measured by earnings); each are listed in the three input screens. These relationships are derived from meta-analytic reviews of the relevant research literature.

For ATOD disorders (including regular smoking), we meta-analyzed two sets of research studies: one set examines the relationship between ATOD disorders and employment rates, and the second examines the relationship between ATOD disorders and earnings, conditional on being employed. Exhibit E2 in Appendix E displays the results of our meta-analysis of these two bodies of research for each ATOD disorder. Our meta-analytic procedures are described elsewhere in this Appendix.

For each ATOD disorder, from these two findings—the effect of ATOD disorders on employment, and the effect of ATOD disorders on the earnings of those employed—we then combined the results to estimate the relationship between an ATOD disorder and average earnings of all people (workers and non workers combined). To do this, we used the effect sizes and standard errors from the meta-analyses on employment and earnings of workers. We used data from the 2009 CPS earnings for average earnings of those with earnings and the standard deviation in those earnings and the proportion of the CPS sample with earnings. We then computed the mean change in earnings for all people by computing the change in the probability of earnings and the drop in earnings for those with earnings. The ratio of total earnings (for both workers and non-workers) for non-disordered individuals to ATOD disordered individuals was then computed.

This mean effect, however, is estimated with error because of the standard errors in the meta-analytic results reported above. Therefore, we used @RISK distribution fitting software to model the joint effects of an alcohol disorder on the mean ratio, given the errors in the two key effect size parameters. The distribution with the best fit (criterion: lowest root-mean squared error) was chosen; for all four disorders, a lognormal distribution was best. Therefore, the two lognormal distribution parameters are entered in the model, as shown in Exhibits D4.a, D4.b, D4.c, and D4.d. Since the body of

evidence we reviewed in the meta analysis did not allow separation of the effects into (1) never disordered people vs. currently disordered people, and (2) formerly disordered people vs. currently disordered people, we enter the same lognormal parameters for both the $EarnGN$ and the $EarnGF$ variables. The sole exception was for smoking, as shown in Exhibit D4.b. Here, the evidence from our review of the literature indicated that former smokers suffer no earnings penalty relative to never smokers. Therefore, we set that parameter to zero.

The present value of the change in morbidity-related earnings for a prevention program that produces a change in the probability of a current ATOD is given by:

$$PV\Delta Earn = \sum_{y=tage}^{65} \frac{(\Delta ATOD_y \times (1 - \sum_{o=1}^y O_o) \times EarnGN \times EarnC_y) + (\Delta ATOD_y \times (1 - (1 - \sum_{o=1}^y O_o)) \times EarnGF \times EarnC_y)}{(1 + dis)^{(y-tage+1)}}$$

Where $\Delta ATOD_y$ is the change in ATOD probability; O_o are the annual onset probabilities; $EarnGN$ is the earnings gain of never-disordered people compared to currently disordered people; $EarnGF$ is the earnings gain of formerly disordered people compared to currently disordered people; dis is the discount rate; and $tage$ is the treatment age of the person in the program. Since a prevention program may serve people without a disorder and with a disorder, the above model weights that probability by the age of onset probabilities.

The present value of the change in the morbidity-related earnings for a treatment program that produces a change in the probability of people with a current ATOD disorder is given by:

$$PV\Delta Earn = \sum_{y=tage}^{65} \frac{(\Delta ATOD_y \times EarnGF \times EarnC_y)}{(1 + dis)^{(y-tage+1)}}$$

This model for a treatment program is simpler than that for a prevention program because, by definition, a treatment program only attempts to turn currently disordered ATOD people into former ATOD people.

$$PVL: P_{mort} = \sum_{a=A}^{100} \frac{PE_a \times R_a \times \sum_{y=a}^{100} (LC_y \times (1 + LGN)) \times DeathPrP_a + PE_a \times (1 - R_a) \times \sum_{y=a}^{100} (LC_y \times (1 + LGF)) \times DeathPrP_a}{(1 + Dis)^a}$$

For labor market morbidity-related benefits for treatment programs, the labor market benefits of ATOD reductions are computed with this equation:

$$PVL: T_{morb} = \sum_{a=A}^{100} \frac{LC_a \times LGF \times PE_a}{(1 + Dis)^a}$$

D4.6 Medical Costs, Treatment Costs, and Other Costs From ATOD

The Institute model computes estimates of changes in avoidable hospital and other medical costs as a result of ATOD morbidity and mortality, including estimates of avoidable treatment costs for alcohol and drug disorders, and for avoidable traffic crash costs for alcohol.

Hospital-Related Parameters. The costs of hospital charges attributable to alcohol, illicit drugs, and smoking, are computed with information from the Washington State Comprehensive Hospital Abstract Reporting System (CHARS) system. CHARS contains hospital inpatient discharge information (derived from billing systems). We use 2007 CHARS data in this analysis. CHARS collects information on billed charges of patients, as well as the codes for their diagnoses. We apply the attributable fraction information, described in D4.3 of this Appendix, to the CHARS data to estimate the number of attributable full time equivalent hospital events by ATOD, $FTEHospitalEvents$, as well as the average billed charge per event, $HospCostEvent$, and the average number of days on an inpatient stay, given a stay. These parameters are shown in Exhibits D4.a, D4.b, and D4.d, for alcohol, tobacco, and illicit drugs, respectively. We also apply a hospital cost-to-charge ratio as described in Appendix D9.

From these inputs, we then compute an upper bound number of events per DSM disorder under the assumption that all classified hospital events stemmed from individuals currently diagnosed with a DSM ATOD disorder (or current regular

smokers for tobacco-related hospital events). A lower bound is calculated assuming that all hospital events stemmed simply from the general use of ATOD, whether or not the use was from DSM disordered populations.

$$ExpHospEventsUpperBound = \frac{FTEHospitalEvents}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y=1}^{100} Pop_y}}$$

$$ExpHospEventsLowerBound = \frac{FTEHospitalEvents}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$ExpHosp\$ = \frac{ExpHospEventUpperBound + ExpHospEventLowerBound}{2} \times HospCostEvent \times CostRatio$$

In computations, the upper bounds and lower bounds form a triangular probability density distribution (with the mean taken as the mode). In Monte Carlo simulation, a random draw is taken from this probability distribution in order to attribute a hospital charge to a disordered DSM ATOD event.

Thus far, the calculations only cover hospitalization costs. Following the work of Rosen et al., we also make an adjustment to include pharmacological drugs and other medical non-durable costs.⁹⁹ To do this, we multiply the expected hospitalization costs, *ExpHosp\$*, by the sum of drug and other non-durable medical costs and total hospital care costs, divided by total hospital care costs. The data for these two cost categories for Washington are the aggregate totals entered in Exhibit D9.a.

Emergency Department Parameters. Emergency Department parameters are shown in Exhibits D4.a for alcohol, D4.b for tobacco, and D4.d for illicit drugs other than cannabis. The model uses a similar approach to that described for hospital events and costs. The model uses an estimate of the probability that an emergency room event is attributable to an alcohol, tobacco, or illicit drug related event. McDonald et al. (2004) estimate 7.9 percent of emergency room visits are alcohol related; Bernstein (2009) estimates 4.9 percent of emergency room visits are tobacco induced; and data from the Drug Abuse Warning Network provide a national estimate of drug-related emergency department visits of 0.84 percent.¹⁰⁰

The total number of emergency department visits in Washington during 2008 is entered in Exhibit D9.a. These data come from a report by the Washington State Hospital Association.¹⁰¹ We then apply the attributable fractions just described; for example, for alcohol, we apply the 7.9 percent causation factor to determine the number of alcohol-related emergency room visits. As with hospital events, we compute upper and lower bound by dividing by the current annual prevalence of DSM disorders in the general population (upper bound) or the current level of use (not just DSM disorders) in the general population (lower bound). We then apply a cost per emergency department event, *EDCostEvent*, and an emergency department cost-to-charge ratio. The cost per emergency department is taken as the median cost from the Medical Expenditure Panel Survey (MEPS) of the U.S. Department of Health & Human Services.¹⁰² In computations, the upper bounds and lower bounds form a triangular probability density distribution (with the mean taken as the mode). In Monte Carlo simulation, a random draw is taken from this probability distribution in order to attribute a emergency department charge to a disordered DSM ATOD event.

$$ExpEDEventsUpperBound = \frac{TotalEDVists \times CausationFraction}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y=1}^{100} Pop_y}}$$

⁹⁹ Rosen et al., 2008

¹⁰⁰ McDonald, A. J., Wang, N., & Camargo, C. A., Jr. (2004). US emergency department visits for alcohol-related diseases and injuries between 1992 and 2000. *Archives of Internal Medicine*, 164(5), 531-537.; Bernstein, S. L. (2009). The clinical impact of health behaviors on emergency department visits. *Academic Emergency Medicine*, 16(11), 1054-1059.; Center for Behavioral Health Statistics and Quality. (2011, February). *Drug abuse warning network, 2008: National estimates of drug-related emergency department visits* (HHS Publication No. SMA 11-4618). Rockville, MD: United States Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Author.

¹⁰¹ Washington State Hospital Association. (2010, October). *Emergency room use* (Developed by WSHA's Health Information Program). Seattle, WA: Author. The Association reports 18 months of data with a total of 2,631,071 visits during the 18 month period from January 2008 to June 2009. We converted this number to an annual estimate for 2008 by multiply by 12/18.

¹⁰² Agency for Healthcare Research and Quality. (2011, June). *Emergency room services-mean and median expenses per person with expense and distribution of expenses by source of payment: United States, 2008* (Medical Expenditure Panel Survey Household Component Data, Table 6). Retrieved June 30, 2011.

$$ExpEDEventsLowerBound = \frac{TotalEDVists \times CausationFraction}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$ExpED\$ = \frac{ExpEDEventsUpperBound + ExpEDEventsLowerBound}{2} \times EDCostEvent \times CostRatio$$

Treatment Parameters. For the cost of admissions for treatment, we undertook an analysis identical to those just described. We obtained the total number of publicly funded treatment events in Washington for alcohol, cannabis, and illicit drugs from the Treatment Episode Data Set (TEDS) of the U.S. Substance Abuse & Mental Health Services Administration. These data are entered in Exhibits D4.a, D4.c, and D4.d. The public cost per treatment is taken from a study of Washington substance abuse treatment by Wickizer in 2007.¹⁰³ We then use the same computational process just described.

$$ExpTreatmentEventsUpperBound = \frac{TotalTreatmentEvents}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y=1}^{100} Pop_y}}$$

$$ExpTreatmentEventsLowerBound = \frac{TotalTreatmentEvents}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$ExpTreatment\$ = \frac{ExpTreatmentEventsUpperBound + ExpTreatmentEventsLowerBound}{2} \times TreatmentCostEvent$$

Traffic Crash Parameters. We modeled alcohol-involved property crash costs with a similar set of procedures. We estimated the annual number of alcohol involved traffic crashes in Washington by obtaining the total number of officer reported traffic collision in Washington in 2009 (102,859).¹⁰⁴ To estimate the proportion of all crashes that are reported by police out of total crashes, we use national estimates produced by Blincoe et al. (2002).¹⁰⁵ Data from Table 3 of Blincoe provide an estimate that 56.7 percent of all crashes are reported by police. Thus, an estimate of total crashes in Washington in 2009 is 181,390. To this we apply the alcohol induced causation factor (8.5 percent) derived from national information also provided in Blincoe et al. (2002), along with the average property crash cost, also from Blincoe et al. (2002) of \$1,891 in 2000 dollars.

$$ExpTrafficCollisionsUpperBound = \frac{TotalTrafficCollisions \times CausationFraction}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y=1}^{100} Pop_y}}$$

$$ExpTrafficCollisionsLowerBound = \frac{TotalTrafficCollisions \times CausationFraction}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$ExpTrafficCollison\$ = \frac{ExpTrafficCollisionsUpperBound + ExpTrafficCollisionsLowerBound}{2} \times TrafficCostEvent$$

D4.7 Age of Initiation of ATOD

As described above, we estimate the costs of disordered use of alcohol, cannabis, other illicit drugs, and regular smoking. These costs are tied to the prevalence of consumption patterns. Many of the ATOD measures used in evaluations of prevention and early intervention programs, however, are measures of the age at initiation of alcohol. Therefore, in order to estimate the long-term costs of disordered ATOD, it is necessary to determine whether there is a causal link between the delay in the age at initiation and the ultimate disordered use of ATOD. For each ATOD disorder, we undertook a review of the literature and contributed original analysis using NSDUH data. Our estimates and sources for these age of initiation parameters are described in Exhibit D4.e.

¹⁰³ Wickizer, T. M. (2007, June). *The economic costs of drug and alcohol abuse in Washington State, 2005*. Olympia: Washington State Department of Social and Health Services, Division of Alcohol and Substance Abuse.

¹⁰⁴ Washington State Department of Transportation. (n.d.). *2009 Washington State collision data summary*. Olympia, WA: Author. Retrieved June 30, 2011 from http://www.wsdot.wa.gov/mapsdata/collision/pdf/Washington_State_Collision_Data_Summary_2009.pdf

¹⁰⁵ Blincoe, L. J., Seay, A. G., Zaloshnja, E., Miller, T. R., Romano, E. O., Luchter, S., & Spicer, R. S. (2002, May). *The economic impact of motor vehicle crashes 2000*. Washington, DC: United States Department of Transportation, National Highway Traffic Safety Administration.

D5. Valuation of Teen Birth Outcomes

In this benefit-cost model, the implications of a teen birth are expressed in terms of the birth's effect on the other outcomes we evaluated. That is, we evaluate the economic consequences of a teen birth based on its relationship to subsequent high school graduation rates, public assistance usage, crime rates, child abuse and neglect cases, K-12 grade repetition, and other outcomes. We evaluate these effects for both the teen mother and the child born to the teen mother. We estimate these effects for births to teens under the age of 18.¹⁰⁶ The results from our meta-analyses of the research literature are shown in Appendix E.

D6. Valuation of Public Assistance Outcomes

Public assistance costs are treated as transfer payments in the benefit-cost model. If a program has an effect on public assistance use, then there is a redistribution of costs between program recipients and taxpayers. For example, if an early childhood education program lowers the use of public assistance, then the reduced public assistance payments are a benefit to taxpayers, but a loss of income to the family in the early childhood assistance program. The only net real cost difference in this transfer is the effect that a change in public assistance caseloads has on costs related to the administration of the public assistance programs.

Exhibit 6.a displays the input screen for this area. Program effects are measured, most often, as a continuous measure of the number of months on public assistance. Therefore, we enter information on Washington State public assistance caseloads including the mean number of months on public assistance for those on the caseload, the standard deviation in the number of months, the average monthly assistance amount, a percentage for agency administrative costs and, for modeling purposes, the age at which public assistance receipt begins.¹⁰⁷

¹⁰⁶ In using the age 18 as a cut-off, we follow the same approach found in Hoffman, S. D., & Maynard, R. A. (Eds.). (2008). *Kids having kids: Economic costs & social consequences of teen pregnancy* (2nd edition). Washington, DC: Urban Institute Press.

¹⁰⁷ The average number of months adults receive TANF in Washington was obtained from : Economic Services Administration. (2009, December). *ESA briefing book: State fiscal year 2009. A reference for programs, caseloads, and expenditures*. Olympia: Washington State Department of Social and Health Services. Retrieved June 30, 2011 from http://www.dshs.wa.gov/pdf/main/briefingbook/2009esa_briefing_book_2009.pdf. The standard deviation was calculated based on a population of female TANF recipients who had participated in an Institute survey in 2008; see: Miller, M. (2011, February). *Depression in Washington's female TANF population: Prevalence, DSHS screening, and treatment* (Document No. 11-02-3401). Olympia: Washington State Institute for Public Policy.

Exhibit 6.a

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

Public Assistance

Back to Main Model

Public Assistance Parameters

| | |
|---|-----|
| Average monthly public assistance benefit | 449 |
| Administrative costs (as a proportion of monthly benefit) | 0.1 |
| Average months on public assistance (lifetime) | 35 |
| Standard deviation (in months) on public assistance | 28 |
| Age at which (for modeling purposes) public assistance receipt begins | 18 |

D7. Model Inputs for K-12 Education Outcomes

D7.1 Input Screens for Education Parameters

Evaluations of education programs or policies often assess outcome measures such as student test scores, years of education, graduation rates, special education, or grade retention. The Institute's benefit-cost model includes a number of education-related parameters to estimate the benefits of these education outcomes. The inputs are entered into the model on a single user screen shown in Exhibit D7.a.

Exhibit D7.a

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

General | **Education** | Economic | Crime

Child Welfare | Substance Use | Health Care | Mental Health | Public Asst | Housing | Teen Birth | Outcomes & Links

Back to Main Model

K-12 Education Parameters

| | All Students | Low Income Students |
|---|--------------|---------------------|
| Gain in lifetime earnings from a 1SD increase in test scores (Mean) | 0.118 | 0.118 |
| Gain in lifetime earnings from a 1SD increase in test scores (Std Error) | 0.03 | 0.03 |
| Gain in lifetime earnings from one extra year of education (Mean) | 0.1 | 0.1 |
| Gain in lifetime earnings from one extra year of education (Std Error) | 0.024 | 0.024 |
| Standard Deviation for number of completed years of education (K-20) | 2.4 | 2.4 |
| High School graduation rate | 0.765 | 0.694 |
| Causal link Between Graduating from High School and Lifetime Earning Gains (Max) | 1 | 1 |
| Causal link Between Graduating from High School and Lifetime Earning Gains (Mode) | 1 | 1 |
| Causal link Between Graduating from High School and Lifetime Earning Gains (Min) | 1 | 1 |
| Retention: Percent retained at least one year in K-12 | 0.098 | 0.165 |
| Retention: Avg number of years retained, for those retained | 1 | 1 |
| Special ed: Percent in special education | 0.126 | 0.156 |
| Special ed: Avg number of years of special ed, for those who receive special ed | 4 | 4 |
| Special ed: Avg age of first entry into special education | 8 | 8 |
| Special ed: Cost of one year of special education per student | 12053 | 12666 |
| Special ed: Year of dollars for cost of special ed | 2010 | 2010 |
| K-12 education: Cost of one year of regular education per student | 7417 | 8030 |
| K-12 education: Year of dollars for cost of regular education | 2010 | 2010 |
| Multiplier for non-market and social benefits of education (Mean) | 0 | 0 |
| Multiplier for non-market and social benefits of education (Std Error) | 0.01 | 0.01 |

The Relationship Between Gains in Student Test Scores and Labor Market Earnings. To evaluate outcomes that measure gains in student standardized test scores, the model contains a parameter and standard error to measure how a one standard deviation gain in test scores relates to a percentage increase in labor market earnings. The standard error for this input is used in Monte Carlo simulations (see Appendix F). Hanushek reviewed the research on this topic and concluded that a one standard deviation gain in math performance in high school is equal to a 12 percent increase in annual earnings.¹⁰⁸ In our own review of the research we found a median 11.8 percent gain in earnings per standard deviation increase in test scores (with a standard error of .03).¹⁰⁹ We enter the same parameter for all students and for low-income students, because our review of the research does not provide separate estimates for low-income populations.

The Relationship Between Gains in Years of Education Completed and Labor Market Earnings. To evaluate outcomes that measure gains in educational attainment, the model contains a parameter and standard error to measure how an extra year of education relates to a percentage increase in labor market earnings. This topic has been one of long-standing interest among economists, and many reviews of the literature are available. For example, Psacharopoulos and Patrinos review many studies from many countries and conclude that “the average rate of return to another year of schooling is 10 percent.”¹¹⁰ Newer estimates employ more rigorous econometric methods to estimate causal effects and have found that returns are usually slightly higher than previous estimates. Heckman et al., however, have found that the estimates vary considerably depending on when the extra year of education occurs. If the extra year leads to high school graduation, for example, the returns are considerably higher than the single point estimates for extra years of college education.¹¹¹ For this reason, we estimate the gains from graduating high school separately, as described below. In our own review of the research, we found a median 10 percent increase in labor market earnings per additional year of education completed (with a standard error of .02).¹¹² We set the same parameter for all students and for low-income students, because our review of the research does not provide separate estimates for low-income populations.

The Standard Deviation in the Number of Completed Years of K-20 education. We used microdata from the March 2009 Current Population Survey to calculate the standard deviation in the number of years of education attained (2.4 years) by adults age 25 or older in the United States.

The High School Graduation Rate. The model contains a user-supplied parameter of the high school graduation rate. The Institute’s entry is Washington State’s on-time graduation rate for 2009-10 published by the Office of Superintendent of Public Instruction.¹¹³ The on-time rate is defined as the percentage of public school students who graduate from high school within four years. This rate is 76.5 percent for all students and 69.4 percent for low-income students.¹¹⁴

¹⁰⁸ Hanushek, E. A. (2009) The economic value of education and cognitive skills. In G. Sykes, B. Schneider, & D. Plank (Eds.), *Handbook of education policy research* (pp. 39-56). New York: Routledge.

¹⁰⁹ We estimated this figure by taking the median of the estimates in Currie, J., & Thomas, D. (2001). Early test scores, school quality and SES: Long-run effects on wage and employment outcomes. In S. Polachek & K. Tatsiramos (Eds.), *Research in labor economics: Vol. 20. Worker wellbeing in a changing labor market* (pp. 103-132). Bingley, UK: Emerald Group; Green, D. A., & Riddell, W. C. (2001). *Literacy, numeracy and labour market outcomes in Canada*. Ottawa, Ontario, Canada: Statistics Canada; Lazear, E. P. (2003). Teacher incentives. *Swedish Economic Policy Review*, 10(2), 179-214; Mulligan, C. B. (1999). Galton versus the human capital approach to inheritance. *Journal of Political Economy*, 107, S184-S224; Murnane, R. J., Willet, J. B., & Levy, F. (1995). The growing importance of cognitive skills in wage determination. *The Review of Economics and Statistics*, 77(2), 251-266; Murnane, R. J., Willett, J. B., Duhaldeborde, Y., & Tyler, J. H. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, 19(4), 547-568.

¹¹⁰ Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, 12(2), 111-134.

¹¹¹ Heckman, J., Lochner, P., & Todd, P. (2008). Earnings functions and rates of return. *Journal of Human Capital*, 2(1), 1-31.

¹¹² We estimated this figure by taking the median of the estimates in Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106(4), 979-1014; Conneely, K., & Uusitalo, R. (1997). *Estimating heterogeneous treatment effects in the Becker schooling model*. Unpublished discussion paper, Industrial Relations Section, Princeton, NJ: Princeton University; Harmon, C., & Walker, I. (1995). Estimates of the economic return to schooling for the United Kingdom. *American Economic Review*, 85(5), 1278-1286; Hausman, J. A., & Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica*, 49(6), 1377-1398; Kane, T., & Rouse, C. E. (1993). *Labor market returns to two- and four-year colleges: Is a credit a credit and do degrees matter?* (NBER Working Paper No. 4268). Cambridge, MA: National Bureau of Economic Research; Maluccio, J. (1997). *Endogeneity of schooling in the wage function*. Unpublished manuscript, Department of Economics, Yale University; Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557-586. These studies are summarized in Card, D. (1999). The causal effect of education on earnings. In E. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (vol. 3, part A, pp. 1801-1863). Atlanta, GA: Elsevier.

¹¹³ Office of Superintendent of Public Instruction. (2011). *Graduation and dropout statistics for Washington in 2009-10*. Olympia: Author. Retrieved June 30, 2011 from <http://www.k12.wa.us/DataAdmin/pubdocs/GradDropout/09-10/GraduationDropoutWashington2009-10.pdf>

¹¹⁴ Low-income students are those eligible for free or reduced-price meals in the National School Lunch Program and School Breakfast Program. Students in households with income up to 130 percent of federal poverty guidelines are eligible for free meals, and those in households up to 185 percent of federal poverty guidelines are eligible for reduced-price meals. For more information visit: <http://www.k12.wa.us/ChildNutrition/Programs/NSLBP/default.aspx>

The Relationship Between High School Graduation and Labor Market Earnings. The model contains a user-supplied parameter to measure the degree of causation between the observed earnings differentials (in the Current Population Survey) for high school graduates and non-graduates. A parameter value of less than one indicates that some of the observed difference is not due, causally, to obtaining a high school diploma but, instead, to other unobserved factors such as motivation or labor market signaling. A zero value implies no causal relationship between any observed differences in earnings, while a value of one indicates that all of the difference in observed earnings is due to the possession of a high school diploma. This parameter is modeled as a triangular probability density distribution. The input screen allows the user to enter a maximum value for this parameter (a value less than or equal to one), a modal value (a value of greater than or equal to zero or less than or equal to one), and a minimum value (a value greater than or equal to zero). The Institute's entries for the maximum, mode, and minimum are set to one. We base these estimates on the work of Rouse¹¹⁵ and Heckman et al.¹¹⁶ Heckman finds very large internal rates of return for high school graduation for both white and black men (they did not study women)—approximately 50 percent. This estimate is in line with an internal rate of return of the difference in earnings observed in the CPS sample (used in this study), given a reasonable up-front level of what Heckman calls “psychic costs” of youths staying in school instead of dropping out.

The K-12 Resource Outcomes. The model can also calculate the value of two other K-12 educational outcomes: years of special education and grade retention. In the user input table shown in Exhibit D7.a, information is entered for the cost of a year of special education, the year in which the special education costs per year are denominated, and the estimated average number of years that special education is used, conditional on entering special education. The user also enters the age when special education is assumed to first be used. The model also requires an estimate of the marginal cost of a year of K-12 education and the year in which these dollars are denominated.¹¹⁷

The Percentage of Students Retained in a Grade Level. The model contains a user-supplied parameter of the percentage of students held back at least one year of school in K-12. The Institute's entry is based on 2009 national rates (9.8 percent of all students and 16.5 percent of low-income students) calculated by the U.S. Department of Education.¹¹⁸ These rates have dropped in recent years; in 1995, 16 percent of U.S. students had been retained in a grade level.¹¹⁹

The Percentage of Students in Special Education. The model contains a user-supplied parameter of the percentage of students in special education. The Institute's entry is the percentage of Washington State students in special education in 2009-10 (12.6 percent).¹²⁰ This rate is not calculated for low-income students in Washington; for this group, we use national estimates of the prevalence of learning disabilities by income level from Planty et al.,¹²¹ to adjust Washington's special education rate to 15.6 percent for low-income students.¹²²

D7.2 Valuation of Earnings From High School Graduation

For any program under consideration that measures high school graduation directly (or indirectly via a “linked” outcome), we use the CPS earnings data and other parameters to estimate the expected gain in life cycle labor market earnings.

First, the annual earnings and benefits are estimated for both high school graduates (*ModEarnHSG*) and non-high school graduates (*ModEarnNHSG*) with the following equations.

¹¹⁵ Rouse, C. E. (2007). Consequences for the labor market. In C. Belfield & H. M. Levin, (Eds.), *The price we pay: Economic and social consequences of inadequate education* (pp. 99-124). Washington, DC: Brookings Institution Press.

¹¹⁶ Heckman et al., 2008.

¹¹⁷ The total cost for one year of special education represents the cost of one year of regular education per student from all sources (state, federal, and local) plus the state allocation for each special education student. The cost of regular education estimate is from: Office of Superintendent of Public Instruction. (2010, March). *Financial reporting summary: School district and educational service district* (Fiscal Year September 1, 2008 – August 31, 2009). Olympia, WA: Author, Table 4. Retrieved June 30, 2011 from <http://www.k12.wa.us/safs/PUB/FIN/0809/0809FinSumweb-7.20.2010.pdf>; the special education allocation estimate is from: Office of Superintendent of Public Instruction. (2011). *OSPI apportionment report for May 31, 2011* (p. 10, report 1220). Retrieved June 30, 2011 from <http://www.k12.wa.us/safs/month.asp>. The average number of years of special education and the average age of first entry in special education are WSIPP estimates.

¹¹⁸ Planty et al. (2009) analyzed the 2003 National Survey of Children's Health and found higher rates of learning disabilities for children in poverty. Planty, M., Hussar, W., Snyder, T., Kena, G., KewalRamani, A., Kemp, J., Bianco, K., & Dinkes, R. (2009). *The condition of education 2009* (NCES 2009-081). Washington, DC: National Center for Education Statistics. Retrieved June 30, 2011 from http://nces.ed.gov/programs/coe/pdf/coe_gra.pdf

¹¹⁹ National Center for Education Statistics. (2006). *The Condition of Education 2006* (NCES 2006-071). Washington, DC: Author. Retrieved June 30, 2011 from http://nces.ed.gov/programs/coe/pdf/coe_grr.pdf

¹²⁰ Office of Superintendent of Public Instruction. *Washington State Report Card*. Retrieved June 30, 2011 from <http://reportcard.ospi.k12.wa.us/summary.aspx?year=2009-10>

¹²¹ Planty et al., 2009.

¹²² We took the percentage of children in special education for up to 185 percent of the federal poverty level divided by the percentage of all children in the United States in special education to determine the factor by which to adjust Washington's special education rate. Altarac, M., & Saroha, E. (2007). Lifetime prevalence of learning disability among US children. *Pediatrics*, 119(Suppl. 1), S77-S83.

$$ModEarnHSG_y = (EarnHSG_y \times (1 + EscHSG)^{y-age}) \times (FHSG \times (1 + EscFHSG)^{y-age}) \times (IPD_{base}/IPD_{cps})$$

$$ModEarnNHSG_y = (EarnNHSG_y \times (1 + EscNHSG)^{y-age}) \times (FNHSG \times (1 + EscFNHSG)^{y-age}) \times (IPD_{base}/IPD_{cps})$$

For each year (y) from the age of a program participant (age) to age 65, the annual CPS earnings for the relevant group (either $EarnHSG$ or $EarnNHSG$, for high school graduates or non-high school graduates) are multiplied by one plus the relevant real earnings escalation rate (either $EscHSG$ or $EscNHSG$) raised to the number of years after program participation, times the relevant fringe benefit rate (either $FHSG$ or $FNHSG$) multiplied by one plus the relevant fringe benefit escalation rate (either $EscFHSG$ or $EscFNHSG$) raised to the number of years after program participation, times a factor to apply the Implicit Price Deflator for the base year dollars, IPD_{base} , chosen for the overall benefit-cost analysis relative to the year in which the CPS data are denominated, IPD_{cps} . In both equations, the two streams of earnings, $EarnHSG$ and $EarnNHSG$, are the annual estimates using the beta probability density distributions discussed above.

The gain in the present value of lifetime earnings from high school graduation is then estimated with this equation:

$$PVEarnGainHSG = \sum_{y=age}^{65} \frac{(ModEarnHSG_y - ModEarnNHSG_y) \times Units_{hsg} \times HSGCC}{(1 + Dis)^{y-age}}$$

For each year from the age of the program participant to age 65, the difference in earnings between high school graduates and non high school graduates is multiplied by the increase in the number of high school graduation “units” (percentage points) caused by the program or policy. The calculation of the units variable was described in Appendices B and C. This product is then multiplied by a parameter to measure the degree of causation ($HSGCC$) between the two present value earnings sums. This last term, which ranges from 0 to 1, can be used if there is evidence that the difference between the two earnings streams ($ModEarnHSG$ and $ModEarnNHSG$) is not due, causally, to obtaining a high school diploma but, instead, to other unobserved factors (such as motivation). A zero value for $HSGCC$ would imply no causal relationship between any observed differences in earnings, while a value of one would indicate that all of the difference in observed earnings is due to the possession of a high school diploma. Sources of estimates for the variable $HSGCC$ are described in Section D.8.1 of this Appendix. The numerator in the equation is then discounted to the age of the program participant (age) with the discount rate (Dis) chosen for the overall benefit-cost analysis.

D7.3 Valuation of Earnings From Increases in K-12 Standardized Student Test Scores

For any program under consideration that measures gains in student standardized test scores directly (or indirectly via a “linked” outcome), we use the CPS earnings data and other parameters to estimate the expected gain in life cycle labor market earnings.

First, the present value of lifetime earnings are estimated for all people, measured with the Current Population Survey with the following equation, where basic CPS earnings are adjusted for long-run real escalation rates and fringe benefit rates and converted into base year dollars. For each year, y , from the age of a program participant, age , to age 65, the modified annual CPS earnings, $ModEarnAll$, are multiplied by one plus the real earnings escalation rate, $EscAll$, raised to the number of years after program participation, times the fringe benefit rate, $FAll$, multiplied by one plus the fringe benefit escalation rate, $EscFAll$, raised to the number of years after program participation, times a factor to apply the Implicit Price Deflator for the base year dollars, IPD_{base} , chosen for the overall benefit-cost analysis relative to the year in which the CPS data are denominated, IPD_{cps} .

$$ModEarnAll_y = (EarnAll_y \times (1 + EscAll)^{y-age}) \times (FAll \times (1 + EscFAll)^{y-age}) \times (IPD_{base}/IPD_{cps})$$

The present value gain in earnings is then estimated. For each year from the age of the program participant to age 65, the modified earnings are multiplied by the increase in the number of test score “units” (standard deviation test score units) caused by the program or policy. The calculation of the units variable is described in Appendices B and C. This term is then multiplied by a parameter to measure the degree of causation, $TSCC$, between a one standard deviation gain in student test scores and the related percentage increase in labor market earnings. Sources of estimates for the variable $TSCC$ are described in Section D.7.1 of this Appendix. The numerator in the equation is then discounted to the age of the program participant, age , with the discount rate, Dis , chosen for the overall benefit-cost analysis.

$$PVEarnGainTS = \sum_{y=age}^{65} \frac{ModEarnAll_y \times Units_{ts} \times TSCC}{(1 + Dis)^{y-age}}$$

D7.4 Valuation of Earnings from Increases in the Number of Years of Education Achieved

For any program under consideration that measures gains in the number of years of education achieved directly (or indirectly via a “linked” outcome), we use the CPS earnings data and other parameters to estimate the expected gain in life cycle labor market earnings.

First, the present value of lifetime earnings are estimated for all people measured with the Current Population Survey with the following equation, where basic CPS earnings are adjusted for long-run real escalation rates and fringe benefit rates and converted into base year dollars. For each year, y , from the age of a program participant, age , to age 65, the modified annual CPS earnings, $ModEarnAll_y$, are multiplied by one plus the real earnings escalation rate, $EscAll$, raised to the number of years after program participation, times the fringe benefit rate, $FALL$, multiplied by one plus the fringe benefit escalation rate $EscFALL$ raised to the number of years after program participation, times a factor to apply the Implicit Price Deflator for the base year dollars, IPD_{base} , chosen for the overall benefit-cost analysis relative to the year in which the CPS data are denominated, IPD_{cps} .

$$ModEarnAll_y = (EarnAll_y \times (1 + EscAll)^{y-age}) \times (FALL \times (1 + EscFALL)^{y-age}) \times (IPD_{base}/IPD_{cps})$$

The present value gain in earnings is then estimated. For each year from the age of the program participant to age 65, the modified earnings are multiplied by the increase in the number of years of education “units” (in standard deviations) caused by the program or policy. The calculation of the units variable is described in Appendices B and C. This term is then multiplied by a parameter to measure the degree of causation, $YearsOfEdCC$, between one extra year of education and the related percentage increase in labor market earnings. Sources of estimates for the variable $YearsOfEdCC$ are described in Section D.7.1 of this Appendix. The numerator in the equation is then discounted to the age of the program participant, age , with the discount rate, Dis , chosen for the overall benefit-cost analysis.

$$PVEarnGainYearsofEd = \sum_{y=age}^{65} \frac{ModEarnAll_y \times Units_{yearsofEd} \times YearsOfEdCC}{(1 + Dis)^{y-age}}$$

D7.5 Valuation of Changes in the Use of K–12 Special Education and Grade Retention

The model can also calculate the value of two other K-12 educational outcomes: years of special education and grade retention. The present value cost of a year of special education is estimated by discounting the cost of a year in special education, $SpecEdCostYear$, for the estimated average number of years that special education is used, conditional on entering special education, $specedyears$. These years are assumed to be consecutive. The present value is to the age when special education is assumed to first be used, $start$. This sum is further present valued to the age of the youth in a program, $progage$, and the cost is expressed in the dollars used for the overall cost benefit analysis, IPD_{base} , relative to the year in which the special education costs per year are denominated, $IPD_{specedcostyear}$.

$$PVspeced_{start} = \sum_{y=1}^{specedyears} \frac{SpecEdCostYear}{(1 + Dis)^y}$$

$$PVspeced_{progage} = \frac{PVspeced_{start} \times \frac{IPD_{base}}{IPD_{specedcostyear}}}{(1 + Dis)^{start-progage}}$$

The present value cost of an extra year of K-12 education is estimated for those retained for an extra year. This is modeled by assuming that the cost of the extra year of K-12 education, $EdCostYear$, after adjusting the dollars to be denominated in the base year dollars used in the overall analysis, would be borne when the youth is approximately 18 years old. Since there is a chance that the youth will not finish high school and, therefore, that the cost of this year will never be incurred, this present valued sum is multiplied by the probability of high school completion, $Hsgradprob$.

$$PVgraderet_{progage} = \left[\frac{EdCostYear \times \frac{IPD_{base}}{IPD_{ed\ cost\ year}}}{(1 + Dis)^{18 - progage}} \right] \times Hsgradprob$$

D7.6 Discount Factors for Decaying Test Score Effect Sizes to Age 17

Many effective education programs increase the standardized test scores of program participants. The magnitude of these early gains, however, does not remain constant over time; researchers have found that test score gains from program participation “fade out” during the K-12 years.¹²³ Our meta-analyses include initial effects size for students’ academic gains on standardized tests relative to the comparison group; a discount factor is then applied to this initial effect size to account for fade-out from the age of measurement to age 17.

We determined the discount factor by performing a multivariate regression analysis of 219 effect sizes spanning the post-test through grade 9 from 47 evaluations of early childhood education programs with multiple follow-up periods. We weighted the model by the inverse variance weight for random effects and included the type of test, type of program, and study research design rating as control variables. The results indicate that by grade 9, test score effect sizes were 41 percent lower than at post-test, on average. We carried these findings out to grade 12 for use in the benefit-cost model. Exhibit D7.b displays the decay rates we used.

Exhibit D7.b

| Age of measurement | Grade | Test score effect size as a percentage of post-test | Fadeout multiplier: Age 17 test score effect size as a percentage of the effect size at age of measurement |
|--------------------|-------|---|--|
| 4 | pre-K | 100% | 47% |
| 5 | K | 96% | 49% |
| 6 | 1 | 92% | 51% |
| 7 | 2 | 88% | 53% |
| 8 | 3 | 84% | 56% |
| 9 | 4 | 79% | 59% |
| 10 | 5 | 75% | 62% |
| 11 | 6 | 71% | 65% |
| 12 | 7 | 67% | 69% |
| 13 | 8 | 63% | 74% |
| 14 | 9 | 59% | 79% |
| 15 | 10 | 55% | 85% |
| 16 | 11 | 51% | 92% |
| 17 | 12 | 47% | 100% |

¹²³ For example, a meta-analysis by Leak et al. (2010) found that early test score gains decreased by at least 54 percent five or more years after the post-test; another meta-analysis by Camilli et al. (2010) estimated that early test score gains fade out by more than 50 percent by age 10; and Goodman & Sianesi (2005) examined fade-out for a single evaluation and found that early test score gains decreased by 30 to 50 percent per follow-up period. Leak, J., Duncan, G., Li, W., Magnuson, K., Schindler, H., & Yoshikawa, H. (2010, November). *Is timing everything? How early childhood education program impacts vary by starting age, program duration, and time since the end of the program*. Paper prepared for presentation at the meeting of the Association for Policy Analysis and Management, Boston, MA; Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, 112(3), 579-620; Goodman, A., & Sianesi, B. (2005). Early education and children's outcomes: How long do the impacts last? *Fiscal Studies*, 26(4), 513-548.

D8. Valuation of Mental Health Outcomes

The Institute's benefit-cost model contains procedures to estimate the monetary value of changes in certain mental health conditions. The model approximates mental health definitions established by the Diagnostic and Statistical Manual (DSM) of the American Psychiatric Association. The current model focuses on ADHD, Depression, Anxiety, and Disruptive Behavior. The latter category covers the DSM categories of Oppositional Defiant Disorder and Conduct Disorder. Obviously, there are other recognized mental health disorders. It is anticipated that future development of the Institute's model will include additional categories. This section of the Technical Appendix describes the Institute's current procedures to estimate the monetary benefits of program-induced changes in these mental health conditions.

In general, the Institute's mental health modeling follows the same analytic procedures described for in Appendix D4 for alcohol, tobacco, and illicit drugs. Readers can refer to that section to find more detail.

The Institute's mental health model uses an incidence-based costing approach. It is not designed to provide an estimate of the total cost to society of current and past mental health disorders. Other studies have attempted to estimate these values.¹²⁴ For example, Insel (2008) summarizes findings indicating the total cost of serious mental illness in the United States in 2002 to be \$317.6 billion in "economic" costs (\$1,081 per capita) with x percent of this total due to health care expenditures, x percent due to loss in labor market earnings, and x percent due to disability payments.¹²⁵ These prevalence-based total cost studies can be interesting, but they are not designed to evaluate future marginal benefits and marginal costs of specific public policy options.

The purpose of the Institute's model is to provide the Washington State legislature with advice on whether there are economically attractive evidence-based policies that, if implemented well, can achieve reductions mental health disorders. To do this, the model monetizes the projected life-cycle costs and benefits of programs or policies that have been shown to achieve improvements—today and in the future—in mental health conditions. If, for example, empirical evidence indicates that a mental health treatment program prevention program can reduce childhood ADHD symptoms, then what long-run benefits, if any, can be expected from this improved outcome? Once computed, the present value of these benefits can be stacked against program costs to determine the relative attractiveness of different approaches to achieve improvements in desired outcomes.

The current version of the mental health model allows the computation of the following types of avoided costs, or benefits, when a program or policy improves the mental health outcomes considered in this model. Depending on each particular mental health disorder, the following benefit or cost categories are included in the Institute's model:

- Labor market earnings from mental health morbidity or mortality, to the degree there is evidence that current earnings are reduced because of mental health disorders (morbidity), or lifetime earnings are lost because of premature death (mortality) caused by mental health disorders.
- Health care costs for mental health morbidity, to the degree that these costs are caused by mental health conditions.
- Value of a statistical life (VSL) estimates, net of labor market gains, applied to the change in mortality (suicide) estimated to be caused by depression.

D8.1 Input Screens for Mental Health Parameters.

The Institute's mental health model is driven with a set of parameters describing various aspects of each disorder's epidemiology and linked relationships with other outcomes. These input parameters are shown on the following four screen shots. In addition, there are several other input parameters used in the mental health model that are general to the Institute's overall benefit-cost model and these are discussed elsewhere in this Appendix. In the following sections, the sources for the parameters and the computational routines are described.

¹²⁴ See, for example, Harwood, H., Ameen, A., Denmead, G., Englert, E., Fountain, D., & Livermore, G. (2000, May). *The economic costs of mental illness, 1992*. Falls Church, VA: The Lewin Group. Retrieved June 30, 2011 from <http://www.lewin.com/content/publications/2487.pdf>; Greenberg, P. E., Kessler, R. C., Birnbaum, H. G., Leong, S. A., Lowe, S. W., Berglund, P. A., & Corey-Lisle, P. K. (2003). The economic burden of depression in the United States: How did it change between 1990 and 2000? *Journal of Clinical Psychiatry*, 64(12), 1465-1475.; Kessler, R. C., Heeringa, S., Lakoma, M. D., Petukhova, M., Rupp, A. E., Schoenbaum, M., . . . Zaslavsky, A. M. (2008). Individual and societal effects of mental disorders on earnings in the United States: Results from the National Comorbidity Survey Replication. *American Journal of Psychiatry*, 165(6), 703-711.

¹²⁵ Insel, T. R. (2008). Assessing the economic costs of serious mental illness. *American Journal of Psychiatry*, 165(6), 663-665.

Exhibits D8.a through D8.d display the parameters for the analysis of mental health disorders.

Exhibit D8.a ADHD

WSIPP Benefit-Cost Model: Version 1.2

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

Mental Health

Back to Main Model

ADHD | Depression | Anxiety | Disruptive Behavior

Attention Deficit Hyperactivity Disorder--Epidemiology

Proportion of general population with lifetime disorder.

Age of Onset of DSM Disorder: the four parameters for a Beta probability density distribution.

Persistence Rate: parameters for a lognormal distribution.

DSM ADHD: Monetary Consequences

Health Care Cost Parameters

Child (ages 1-18) annual additional health care for ADHD. Year of dollars

Adult (>18) annual additional health care cost for ADHD. Year of dollars

Exhibit D8.b Depression

WSIPP Benefit-Cost Model: Version 1.2

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

Mental Health

Back to Main Model

ADHD | **Depression** | Anxiety | Disruptive Behavior

Depression--Epidemiology

Proportion of general population with lifetime disorder:

Age of Onset of DSM Disorder: the four parameters for a Beta probability density distribution.

Persistence Rate: parameters for a beta distribution.

DSM Depression: Monetary Consequences

Labor Market parameters

| | Mean | Std dev |
|---|------------------------------------|------------------------------------|
| Gain in labor market earnings for never disordered vs current disorder, lognormal probability density distribution parameters | <input type="text" value="0.096"/> | <input type="text" value="0.034"/> |
| Gain in labor market earnings for former disorder vs current disorder, lognormal probability density distribution parameters | <input type="text" value="0.096"/> | <input type="text" value="0.034"/> |

Health Care Cost Parameters

Child (ages 1-18) annual additional health care for depression. Year of dollars

Adult (>18) annual additional health care cost for depression. Year of dollars

Annual Depression Attributed Deaths

| Lower bound of age group | Upper bound of age group | Number of suicides in state (all causes) | Proportion of suicides attributable to DSM disorder | State deaths (all) | State population in age group |
|--------------------------|--------------------------|--|---|--------------------|-------------------------------|
| 1 | 14 | 3.6 | 0.5 | 615.6 | 1251485 |
| 15 | 19 | 42 | 0.5 | 243.2 | 432244.2 |
| 20 | 24 | 64 | 0.5 | 355.6 | 434752 |
| 25 | 34 | 117.2 | 0.5 | 713.4 | 869927.6 |
| 35 | 44 | 160.4 | 0.5 | 1453 | 939210.8 |

Average annual data over the period:

Exhibit D8.c Anxiety

WSIPP Benefit-Cost Model: Version 1.2 X

Enter Sector Inputs
Enter Program Inputs
Run Models & View Reports

General

Economic

Crime

Education

Child Welfare

Substance Use

Health Care

Mental Health

Public Asst

Housing

Teen Birth

Outcomes & Links

Mental Health X

Back to Main Model

ADHD
Depression
Anxiety
Disruptive Behavior

Anxiety--Epidemiology

Proportion of general population with lifetime disorder. 0.315

Age of Onset of DSM Disorder: the four parameters for a Beta probability density distribution. 0.40667

2.1615

5

79

Persistence Rate: parameters for a beta distribution. 0.82942

2.0051

0

196.67

DSM Anxiety: Monetary Consequences

Labor Market parameters

| | Mean | Std dev |
|---|-------|---------|
| Gain in labor market earnings for never disordered vs current disorder, lognormal probability density distribution parameters | 0.215 | 0.171 |
| Gain in labor market earnings for former disorder vs current disorder, lognormal probability density distribution parameters | 0.215 | 0.171 |

Health Care Cost Parameters

| | | | | |
|--|--|------|------|-----------------|
| | Child (ages 1-18) annual additional health care for anxiety. | 1599 | 2007 | Year of dollars |
| | Adult (>=18) annual additional health care cost for anxiety. | 3509 | 2007 | Year of dollars |

91

Exhibit D8.d Disruptive Behaviors

WSIPP Benefit-Cost Model: Version 1.2 X

Enter Sector Inputs
Enter Program Inputs
Run Models & View Reports

General

Economic

Crime

Education

Child Welfare

Substance Use

Health Care

Mental Health

Public Asst

Housing

Teen Birth

Outcomes & Links

Mental Health X

[Back to Main Model](#)

ADHD | Depression | Anxiety | **Disruptive Behavior**

Disruptive Behavior Disorders--Epidemiology

Proportion of general population with lifetime disorder. 0.09

Age of Onset of DSM Disorder: the four parameters for a Beta probability density distribution.

1.8705 1.2511

3 18

Persistence Rate: parameters for a lognormal distribution.

1.869 1.122

DSM Disruptive Behavior: Monetary Consequences

Health Care Cost Parameters

1924
Child (ages 1-18) annual additional health care cost for CD/ODD.
2007
Year of dollars

1924
Adult (>18) annual additional health care cost for CD/ODD.
2007
Year of dollars

Exhibit D8.e
Input Parameters for the Epidemiology of Mental Health Disorders

| | DSM ADHD | DSM Depression | DSM Anxiety | Disruptive Behavior |
|---|--------------|----------------|--------------|---------------------|
| | (a) | (b) | (c) | (d) |
| Percent of population with lifetime DSM disorder ⁽¹⁾ | 8.1% | 23.2% | 31.5% | 9.0% |
| Age of onset | | | | |
| Type of distribution ⁽²⁾ | Beta-general | Beta-general | Beta-general | Beta-general |
| Parameter 1 | 17.362 | 1.1615 | .40667 | 1.8705 |
| Parameter 2 | 41.582 | 2.1852 | 2.1615 | 1.2511 |
| Parameter 3 | 3 | 9 | 5 | 3 |
| Parameter 4 | 18 | 79 | 79 | 18 |
| Persistence of DSM disorder, given onset | | | | |
| Type of distribution ⁽³⁾ | Lognormal | Beta-general | Beta-general | Loglogistic |
| Parameter 1 | 3.2391 | .51946 | .82942 | 0 |
| Parameter 2 | 1.5097 | 2.6936 | 2.0051 | 6.5365 |
| Parameter 3 | n/a | 0 | 0 | 1.537 |
| Parameter 4 | n/a | 138.09 | 196.67 | n/a |

Notes and sources

1. . Kessler, R.C., Berglund, P., Delmer, O., Jin, R., Merikangas, K.R., & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry, 62(6): 593-602. Estimates from Table 3; the estimate for disruptive behavior is an average of the reported risk for oppositional-defiant disorder and conduct disorder.
2. All age of onset distributions were fit with data reported in Kessler, R.C., Berglund, P., Delmer, O., Jin, R., Merikangas, K.R., & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry, 62(6): 593-602. From Table 3 in the paper, we estimated probability density distributions for the age of onset of each of the four mental health disorders, conditional on having a disorder. @Risk software was used to estimate alternative distributions; the distribution with the best fit (criterion: lowest root-mean squared error) was chosen. Beta-general distributions were the best fitting.
3. To estimate persistence of DSM mental health disorders we used the publicly available information from the National Comorbidity Survey-Replication (NCS-R). The NCS-R surveyed a representative sample of 9,282 adults in the United States in 2001-03 to estimate prevalence of mental illnesses in the U.S. population. We identified persons with a lifetime diagnosis of attention deficit, behavioral, any anxiety major depressive disorders. For each disorder we calculated the interval from first to last episode. Those without an episode in the prior 12 months were considered to be free of the disorder. For each disorder, we used survival analysis and the appropriate survey weight to model time to remission. We then used these data to fit the parameters of probability distributions that fit the data. @Risk software was used to estimate alternative distributions; the distribution with the best fit (criterion: lowest root-mean squared error) was chosen, and the winning distribution, and its parameters, is shown for each mental health disorder.

D8.2 Mental Health Epidemiological Parameters

The Institute's mental health model begins by analyzing the epidemiology of each mental health disorder to produce estimates of the current 12-month prevalence. An estimate of the current prevalence of each disorder is central to the benefit-cost model because, for dichotomously measured outcomes, it becomes the "base rate" to which program or policy effect sizes are applied to calculate the change in the number of avoided mental health "units" caused by the program, over the lifetime following treatment.

Four parameters enter the model to enable an estimate of the current prevalence of each mental health disorder, from age 1 to age 100.

- **Lifetime prevalence:** the percentage of the population that has a specific lifetime mental health disorder.
- **Age of onset:** the age of onset of the specific mental health disorder.
- **Persistence:** the persistence of the specific mental health, given onset.
- **Death (Survival):** the probability of death by age, after the age of treatment by a program.

The parameters that enter the model appear on each screen shot on Exhibits D8.a through D8.d. Exhibit D8.e also displays the current parameters in the Institute's model for the first three epidemiological factors, along with sources and notes. The death probability information is described elsewhere in this Appendix.

For each mental health disorder, the current prevalence of the disorder is estimated with this equation.

$$(D.8.1) \quad CP_y = \left(\sum_{0=1}^y O_0 \times P_{(y-0+1)} \right) \times LTP \times S_y$$

The current disorder prevalence probability at any year in a person's life, CP_y , is computed with information on the age-of-onset probability, O_y , from prior ages to the current age of the person, times the persistence probability, P , of remaining in the DSM condition at each onset age until the person is the current age, times the lifetime probability of ever having the DSM disorder, LTP , times the probability of survival at each age, S_y , following treatment by a program.

For each mental health disorder, the exogenous age-of-onset probability distribution for ages 1 to 100, O_y , is a density distribution and is estimated with information from the sources shown in Exhibit D8.e. The parameters in Exhibit D8.e are the same as those entered by the user on the screen shots in Exhibits D8.a through D8.d.

$$1 = \sum_{y=1}^{100} O_y$$

Also, for each mental health disorder, the exogenous persistence distribution for ages after onset, P , is computed from the sources shown in Exhibit D8.e. The persistence distribution describes the probability, on average, of being in the DSM disorder condition each year following onset.

The probability of survival at any given age, S_y , is computed from a national life table on survival, LTS , in the general population. The inputs for the survival table are described in another section of this Technical Appendix. To compute the current prevalence of a disorder over the entire life course, S_y is normalized to age 1.

$$S_y = \frac{LTS_y}{LTS_1}$$

Since the probability of survival depends on the number still living at the treatment age, $tage$, the S_y is normalized to the age of the person being treated in the program being analyzed, since it is assumed that all treatment programs will be for those currently alive at time of treatment.

$$S_y = \frac{LTS_y}{LTS_{tage}}$$

Equation D.8.1 describes the calculation of current prevalence for general (prevention) populations. For programs treating indicated populations, CP_y in equation D.8.2 describes the prevalence in all years following treatment.

$$(D.8.2) \quad CP_y = \frac{\sum_{0=1}^{tage} O_0 \times P_{(y-0+1)}}{\sum_{0=1}^{tage} O_0} \times S_y \times SF$$

The additional term in equation D.8.2 is the reduced chance of survival for someone with depression. We compute an estimate for this as a single parameter with the following equation.

$$SF = \frac{\sum_{a=1}^A \left(Pop_a \times CP_a \times \frac{(PopD_a - DeprD_a)}{\frac{Pop_a}{PopD_a}} \right)}{\sum_{a=1}^A (Pop_a \times CP_a)}$$

In this equation, Pop_a is the total population in a state in each age group, CP_a is the average current depression disorder prevalence in each age group, $PopD_a$ is the total number of deaths in a state in each age group, and $DeprD_a$ is the deaths attributable to depression –induced suicides in each age group. The suicide data are entered on Exhibit D8.b. The suicide death data are obtained from the Washington State Department of Health.

D8.3 Linkages: Mental Health to Other Outcomes

The Institute's benefit-cost model monetizes improvements in mental health outcomes, in part, with linkages between each mental health outcome and other outcomes to which a monetary value can be estimated. The parameters for these linkages are obtained by a meta-analytic review of relevant research literature. For example, we estimate the relationship between DSM mental health conditions and labor market earnings by meta-analyzing the most credible studies that have addressed this topic. The meta-analytic process provides both an expected value effect given the weight of the evidence, and an estimate of the error of the estimated effect. Both of these two parameters are entered into the benefit-cost model and used when performing a Monte Carlo simulation. The linkages in the current Institute model are listed in Appendix E.

D8.4 Human Capital Outcomes Affecting Labor Market Earnings via Mental Health Morbidity and Mortality

The Institute model computes lost labor market earnings, as a result of mental health morbidity and mortality, when there is evidence that the linkage is causal. The procedures begin by estimating the labor market earnings of an average person with a current DSM mental health disorder. As described in Appendix D.1, the Institute's model uses national earnings data from the U.S. Census Bureau's Current Population Survey. The CPS data used in this analysis represent average earnings of all people, both workers and non-workers at each age.

For each person at each age, total CPS earnings can be viewed as a weighted sum of people who have never had a mental health disorder, plus those that are currently disordered, plus those that were formerly disordered. From the CPS data on total earnings for all people, the earnings of individuals with a current mental health condition, at each age, y , is computed with this equation:

$$EarnC_y = \frac{EarnAll_y \times (1 + EarnEscAll)^{y-tage} \times EarnBenAll \times (1 + EarnBenEscAll)^{y-tage} \times (IPD_{base}/IPD_{cps})}{\left((1 + EarnGN) \times \left(1 - (CP_y + (\sum_{0=1}^y (O_0 \times LTP) - CP_y)) \right) + (1 + EarnGF) \times (\sum_{0=1}^y (O_0 \times LTP) - CP_y) + CP_y \right)}$$

The numerator in the above equation includes the CPS earnings data for all people, $EarnAll$, with adjustments for real earnings growth, $EarnEscAll$, earnings-related benefits, $EarnBenAll$, growth rates in earnings benefits, $EarnBenEscAll$, and an adjustment to denominate the year of the CPS earnings data, IPD_{cps} , with the year chosen for the overall analysis, IPD_{base} . These variables are described in Appendix D.1.

The denominator uses the epidemiological variables described above: age of onset probabilities, O_y , lifetime prevalence rates, LTP , and current 12-month prevalence rates, CP_y , at each age.

The denominator also includes two variables on the earnings gain of never-disordered people compared to currently disordered people, $EarnGN$, and the earnings gain of formerly disordered people compared to currently disordered people, $EarnGF$. These two central relationships measure the effect of a DSM mental health condition on labor market

success (as measured by earnings). These relationships are derived from meta-analytic reviews of the relevant research literature as listed in Appendix E.

For mental health disorders, we meta-analyzed two sets of research studies: one set examines the relationship between mental health disorders and employment rates, and the second examines the relationship between mental health disorders and earnings, conditional on being employed. Exhibit E2 in Appendix E displays the results of our meta-analysis of these two bodies of research for DSM mental health disorders. Our meta-analytic procedures are described elsewhere in this Appendix.

For a mental health disorder, from these two findings—the effect of a mental health disorder on employment, and the effect of a mental health disorder on the earnings of those employed—we then combine the results to estimate the relationship between a mental health disorder and average earnings of all people (workers and non workers combined). To do this, we use the effect sizes and standard errors from the meta-analyses on employment and earnings of workers. We use data from the 2009 CPS earnings for average earnings of those with earnings and the standard deviation in those earnings and the proportion of the CPS sample with earnings. We then compute the mean change in earnings for all people by computing the change in the probability of earnings and the drop in earnings for those with earnings. The ratio of total earnings (for both workers and non-workers) for non-disordered individuals to mental health disordered individuals is then computed.

This mean effect, however, is estimated with error as measured by the standard errors in the meta-analytic results reported above. Therefore, we use @RISK distribution fitting software to model the joint effects of an alcohol disorder on the mean ratio, given the errors in the two key effect size parameters. The distribution with the best fit (criterion: lowest root-mean squared error) is a lognormal distribution. Therefore, the two lognormal distribution parameters are entered in the model, as shown in Exhibits D4.b, D4.c. Since the body of evidence we reviewed in the meta analysis did not allow separation of the effects into (1) never disordered people vs. currently disordered people, and (2) formerly disordered people vs. currently disordered people, we enter the same lognormal parameters for both the *EarnGN* and the *EarnGF* variables.

The present value of the change in morbidity-related earnings for a prevention program that produces a change in the probability of a current mental health disorder is given by:

$$PV\Delta Earn = \sum_{y=tage}^{65} \frac{(\Delta MH_y \times (1 - \sum_{o=1}^y O_o) \times EarnGN \times EarnC_y) + (\Delta MH_y \times (1 - (1 - \sum_{o=1}^y O_o)) \times EarnGF \times EarnC_y)}{(1 + dis)^{(y-tage+1)}}$$

Where ΔMH_y is the change in mental health disorder probability; O_o are the annual onset probabilities; *EarnGN* is the earnings gain of never-disordered people compared to currently disordered people; *EarnGF* is the earnings gain of formerly disordered people compared to currently disordered people; *dis* is the discount rate; and *tage* is the treatment age of the person in the program. Since a prevention program may serve people without a disorder and with a disorder, the above model weights that probability by the age of onset probabilities.

The present value of the change in the morbidity-related earnings for a treatment program that produces a change in the probability of people with a current mental health disorder is given by:

$$PV\Delta Earn = \sum_{y=tage}^{65} \frac{(\Delta MH_y \times EarnGF \times EarnC_y)}{(1 + dis)^{(y-tage+1)}}$$

This model for a treatment program is simpler than that for a prevention program because, by definition, a treatment program only attempts to turn currently disordered people into formerly disordered people.

$$PVL: P_{mort} = \sum_{a=A}^{100} \frac{PE_a \times R_a \times \sum_{y=a}^{100} (LC_y \times (1 + LGN)) \times DeathPrP_a + PE_a \times (1 - R_a) \times \sum_{y=a}^{100} (LC_y \times (1 + LGF)) \times DeathPrP_a}{(1 + Dis)^a}$$

For labor market morbidity-related benefits for treatment programs, the labor market benefits of mental health disorder reductions are computed with this equation:

$$PVL: T_{morb} = \sum_{a=A}^{100} \frac{LC_a \times LGF \times PE_a}{(1 + Dis)^a}$$

D8.5 Medical Costs

The Institute's model computes health care costs incurred (or avoided) with changes in the mental health conditions modeled. The inputs for these parameters are shown on Exhibits D8.a through D8.d. They were computed from an analysis of data from the federal Medical Expenditure Panel Survey (MEPS).

Estimates for Mental Disorders:

The Medical Expenditure Panel Survey (MEPS) is a nationally representative large-scale survey of American families, medical providers, and employers who report on healthcare service utilization and associated medical conditions, costs, and payments. An annual cost of healthcare services for 2007 (the latest available year of the MEPS) was calculated by adding the costs of inpatient, outpatient, emergency room, home health, and prescription medication costs per individual. This figure was regressed on a dummy variable representing a mental disorder of interest, controlling for demographic variables, psychiatric comorbidity, and other factors that might be expected to simultaneously correlate with mental illness and inflate total healthcare costs (e.g., existence of chronic illnesses, child delivery, health insurance). The resulting regression coefficient for each disorder represents an estimate of the additional cost of healthcare service utilization per year to individuals with the disorder versus without it. Separate regression models were conducted for adults (over 18 years old) and children, because the coefficients for some disorders were different by age groups.

The costs described above were modified in several ways: First, in order to compute costs that are reflective of the present time, figures from the 2007 MEPS were adjusted for inflation and escalation in healthcare costs over time. Second, we were concerned that for some disorders the psychosocial interventions we reviewed would not substitute for medication use (i.e., that medications would continue to be a cost for individuals with this disorder even after successful psychosocial treatment). If this were the case, our figures – based on a total annual cost that includes prescription medications - would overestimate the benefits that would be expected from an effective intervention. Thus, for several disorders (ADHD, Bipolar Disorder, Schizophrenia), medication costs were removed from the total annual costs, such that the additional costs attributed to these disorders are only for inpatient, outpatient, emergency room, and home health services. Lastly, some disorders (e.g., Conduct and Oppositional Defiant Disorders) are reported by so few MEPS respondents that we were concerned about the representativeness of these individuals relative to the larger population. In such cases, we applied figures from analyses of other disorders that we believed to be similar (for instance, figures for CD/ODD are derived from analyses of ADHD).

D9. Health Care Parameters

The benefit-cost model uses a number of health care parameters. These are shown on the screen shot in Exhibit D9.a.

Total Washington personal health care expenditures are collected for 2004, the most recent year available from the Centers for Medicare & Medicaid Services, U.S. Department of Health & Human Services.¹²⁶ Information on who pays for personal health care expenditures is from the same source, but uses more recent 2009 national data.¹²⁷

Exhibit D9.a

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs Enter Program Inputs Run Models & View Reports

General **Health Care**

Economic [Back to Main Model](#)

Crime

Education

Child Welfare

Substance Use

Health Care

Mental Health

Public Asst

Housing

Teen Birth

Outcomes & Links

State Personal Health Care Expenditures

| Category | Amount |
|--------------------------------------|------------------|
| Total | \$31,600,000,000 |
| Hospital care | \$10,702,000,000 |
| Physician and clinical services | \$9,004,000,000 |
| Other professional services | \$1,363,000,000 |
| Dental services | \$2,505,000,000 |
| Home health care | \$823,000,000 |
| Drugs and other medical non-durables | \$3,792,000,000 |
| Durable medical products | \$485,000,000 |
| Nursing home care | \$1,860,000,000 |
| Other personal health care | \$1,065,000,000 |

Amount:

Year of Data:

Personal Health Care Expenditures by Payer (Percent)

| | |
|--|-------|
| Participant (out of pocket) | 0.143 |
| Taxpayer (medicare and medicaid) | 0.432 |
| Other (including private insurance premia) | 0.425 |

Long-run real escalation rate in health care costs

| | |
|------|-------|
| mode | 0.018 |
| low | 0.005 |
| high | 0.027 |

Avg hospital cost to charge ratio

Emergency Department-related parameters, annual

Total ED admissions: Year of data:

ED cost to charge ratio:

| | |
|--|-------|
| Participant (out of pocket) | 0.174 |
| Taxpayer (medicare and medicaid) | 0.426 |
| Other (including private insurance premia) | 0.396 |

Average Medical Costs, by educational attainment

| Age | Less than high school graduate | | | At least high school graduate | | |
|-----|--------------------------------|--------|-----------|-------------------------------|--------|-----------|
| | Personal | Public | Insurance | Personal | Public | Insurance |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 102 | 440 | 522 | 82 | 225 | 166 |

Year of dollars for average medical costs:

| | |
|--|-------|
| Personal (out of pocket) causal factor for relationship between high school grad and health care costs | 0.72 |
| Taxpayer causal factor for relationship between high school grad and health care costs | 0.9 |
| Private Insurance causal factor for relationship between high school grad and health care costs | 0.74 |
| Odd Ratio: mean survival probability for high school graduates vs. general population survival | 0.447 |

¹²⁶ Centers for Medicare & Medicaid Services, *Health expenditures by state of residence, 1991-2004*. Retrieved June 30, 2011 from http://www.cms.gov/NationalHealthExpendData/05_NationalHealthAccountsStateHealthAccountsResidence.asp#TopOfPage

¹²⁷ Centers for Medicare & Medicaid Services, U.S. Department of Health & Human Services. Retrieved June 30, 2011 from <http://www.cms.gov/NationalHealthExpendData/downloads/tables.pdf>, Table 6, data for 2009.

A hospital cost-to-charge ratio for Washington State is computed with 2009 data from the Healthcare Cost and Utilization Project (HCUP) of the U.S. Department of Health & Human Services.¹²⁸

An estimate of the long-run real escalation rate in per capital inflation-adjusted personal health care costs is computed from the 2009-2019 forecast from Centers for Medicare & Medicaid Services, U.S. Department of Health & Human Services.¹²⁹

Total annual emergency room visits in Washington for 2008 is computed from data compiled by the Washington State Hospital Association.¹³⁰ Information on emergency room charges by type of payer are obtained from the Agency for Healthcare Research and Quality, U.S. Department of Health & Human Services.¹³¹

Health Care cost estimates for high school graduation compared to less than high school graduation:

The Medical Expenditure Panel Survey (MEPS) is a nationally representative large-scale survey of American families, medical providers, and employers who report on healthcare service utilization and associated medical conditions, costs, and payments. An annual cost of services paid by public (e.g., Medicaid, Medicare), private (i.e., insurance), and personal (i.e., family out-of-pocket) sources was computed for 2007 (the latest available year of the MEPS). Among adults, mean costs were analyzed by age and high school graduation status (whether the individual has at least a high school diploma), such that at each age a difference between those with and without a diploma in public, private, and personal costs could be computed.

These mean differences are descriptive in nature and do not account for demographic or other differences between individuals with and without a high school diploma that could influence healthcare costs. As such, ordinary least squares (OLS) regression models were conducted for each payment source. First, we analyzed a sparsely controlled model, which produced a regression coefficient for the additional cost of having a diploma that was similar to the descriptive data. Next, we analyzed a regression model that included multiple covariates for demographic variables as well as other factors that might be expected to simultaneously correlate with education and inflate total healthcare costs (e.g., childbirth). As expected, the regression coefficient for having a high school diploma in the highly controlled model was smaller than the sparsely controlled model. The difference between the estimates from these two models is reflected in the “causal factor” listed for each payment source (public, private, personal) in the screen shot. For example, the difference in personal healthcare costs between individuals with and without a high school diploma is not the mean difference displayed in the blue table, but that difference multiplied by 0.72.

¹²⁸ Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project: <http://hcupnet.ahrq.gov/>

¹²⁹ Centers for Medicare & Medicaid Services. (n.d.). *National health expenditure projections 2009-2019*. United States Department of Health & Human Services, Author. Retrieved June 30, 2011 from <http://www.cms.gov/NationalHealthExpendData/downloads/proj2009.pdf>

¹³⁰ Washington State Hospital Association. (2010, October). *Emergency room use*. Seattle, WA: Author. The table on page 4 reports 18 months of emergency department visits for January 2008 to June 2009. This sum was multiplied by 2/3 to convert to an annual figure representing the year 2008.

¹³¹ Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project: <http://hcupnet.ahrq.gov/>

D10. Other Parameters

In addition to the parameters discussed in the previous sections of this Appendix, the model uses a number of additional user-supplied inputs to compute benefits and costs. These are discussed in this section.

D10.1 Base Year for Monetary Denomination

The model contains many price and monetary values; each is denominated in a particular year's monetary values. To express all monetary values in a common year, the user selects a base year. When the model runs, all monetary values entered into the model are converted to the base year values with the price index chosen by the user (see Section D10.4). The input screen for the base year is shown in Exhibit D.10a.

Exhibit D10.a

The screenshot displays the 'WSIPP Benefit-Cost Model: Version 1.1' window. At the top, there are three tabs: 'Enter Sector Inputs', 'Enter Program Inputs', and 'Run Models & View Reports'. Below these is a 'General' sub-window with a 'Back to Main Model' button. The main area is divided into five sections: 'Base Year for Dollars', 'Discount Rates', 'Demographic', 'VSL', and 'Deadweight Cost'. The 'Base Year for Dollars' section contains a text input field with the value '2010' highlighted in yellow. On the left side of the window, there is a vertical menu of buttons for various sectors: 'General', 'Economic', 'Crime', 'Education', 'Child Welfare', 'Substance Use', 'Health Care', 'Mental Health', 'Public Asst', 'Housing', 'Teen Birth', and 'Outcomes & Links'.

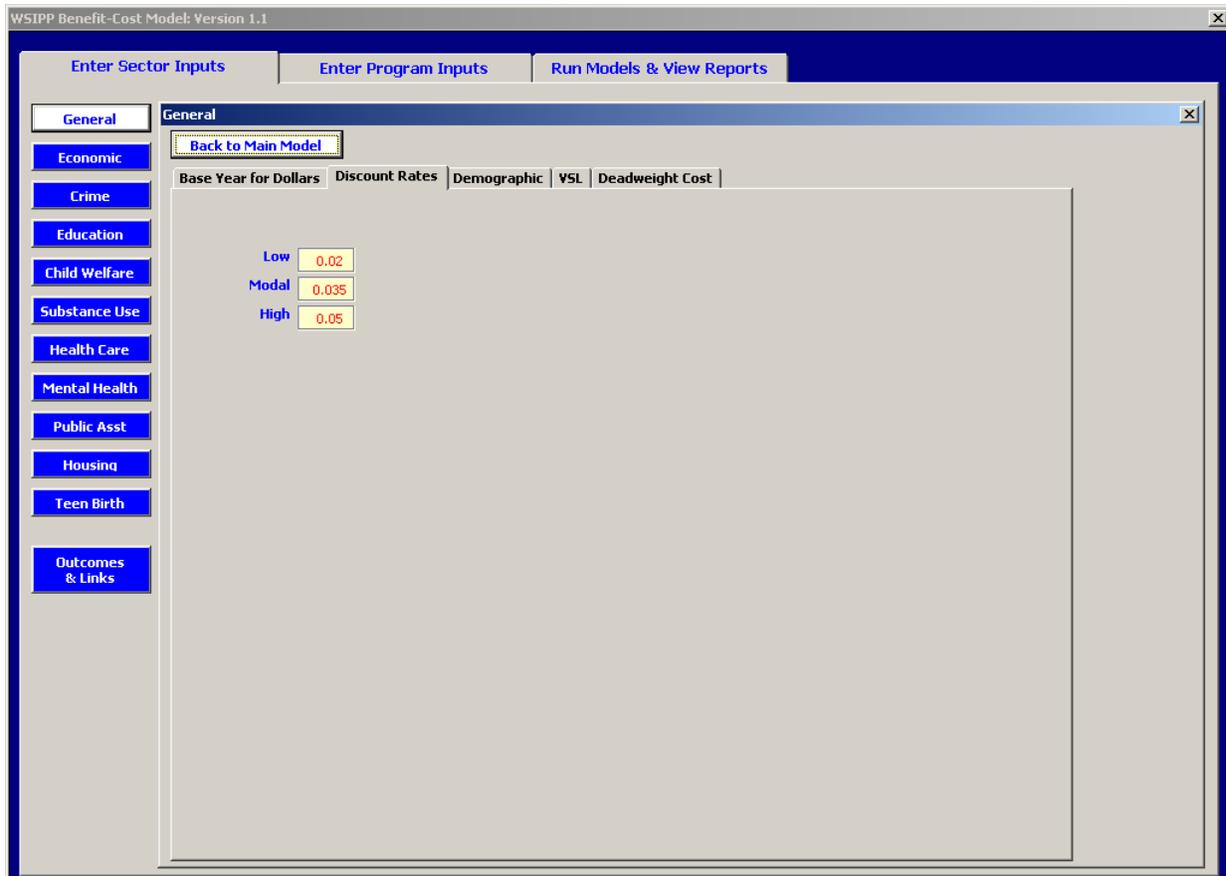
D10.2 Discount Rates

The model uses a range of real discount rates to compute net present values. The discount rates are applied to all annual benefit and cost cash flows and presented-valued to the time the investment would be made. Equation D10.1 indicates that the net present value of a program, evaluated at the at the age of a person for whom an investment is made, NPV_{age} , is the discounted sum of benefits at each year, B_y , minus program costs at each year, C_y , discounted with a discount rate, Dis .

$$(D10.1) \quad NPV_{age} = \sum_{y=age}^N \frac{B_y - C_y}{(1 + Dis)^y}$$

The model uses low, modal, and high discount rates in computations. When the model is run in non-simulation mode, the modal discount rate is used. In Monte Carlo simulation, each run randomly draws a discount rate from a triangular probability density distribution, with the user-selected low, modal, and high discount rates defining the triangle. Exhibit D10.b is a screen shot showing where the three discount rates are entered. The Institute uses a low real discount rate of 2 percent, a modal rate of 3.5 percent, and a high rate of 5 percent. These input choices reflect the recommended rates in Moore et al. (2004).¹³² Similarly, the Congressional Budget Office has used a 3 percent real discount rate in its analyses of Social Security.¹³³ Heckman et al. (2010) analyzed the benefits and costs of the Perry Preschool program and employed a range of discount rates; they used a 3 percent rate to summarize the main benefit-cost results.¹³⁴

Exhibit D10.b



The screenshot shows the 'WSIPP Benefit-Cost Model: Version 1.1' interface. The 'Enter Program Inputs' tab is active, and the 'Discount Rates' sub-tab is selected. The 'Low' rate is set to 0.02, the 'Modal' rate is 0.035, and the 'High' rate is 0.05. A 'Back to Main Model' button is visible at the top of the input area.

| Discount Rate Type | Value |
|--------------------|-------|
| Low | 0.02 |
| Modal | 0.035 |
| High | 0.05 |

¹³² Moore, M. A., Boardman, A. E., Vining, A. R., Weimer, D. L., & Greenberg, D. H. (2004). "Just give me a number!" Practical values for the social discount rate. *Journal of Policy Analysis and Management*, 23(4), 789-812.

¹³³ Congressional Budget Office. (2006, June). *Updated long-term projections for social security*. Washington, DC: Author. Retrieved June 30, 2011 from <http://www.cbo.gov/ftpdocs/72xx/doc7289/06-14-LongTermProjections.pdf>

¹³⁴ Heckman et al., 2010.

D10.3 Demographic Information

Several of the computations in the model require basic demographic information about population in the jurisdiction to which the model is applied. Exhibit D10.c displays the screen shot for these inputs. The total annual population for the jurisdiction by year is included along with a forecast. For Washington State, we enter the total state population estimates from the Washington State Office of Financial Management (OFM), the official forecasting agency for the state. The model also requires information on the current distribution of the state population by single year of age. For Washington, we enter this information as supplied by OFM. Finally, the model needs a recent life table with information on the number of people in a birth cohort surviving to each year along with the life expectancy. We use life table information for the United States produced by the U.S. Department of Health and Human Services Centers for Disease Control and Prevention.¹³⁵

Exhibit D10.c

The screenshot shows the 'General' input screen for the 'Demographic' section. It features three data tables and a 'Year of Cohort' input field.

| Year | Number |
|------|---------|
| 1970 | 3413244 |
| 1971 | 3436299 |
| 1972 | 3430299 |
| 1973 | 3444299 |
| 1974 | 3508700 |
| 1975 | 3567901 |
| 1976 | 3634904 |
| 1977 | 3715400 |
| 1978 | 3836199 |
| 1979 | 3979199 |
| 1980 | 4132156 |
| 1981 | 4229278 |
| 1982 | 4276549 |
| 1983 | 4307247 |
| 1984 | 4354067 |

| Age | Number |
|-----|--------|
| 1 | 87204 |
| 2 | 83618 |
| 3 | 84017 |
| 4 | 84483 |
| 5 | 83334 |
| 6 | 83927 |
| 7 | 85798 |
| 8 | 83821 |
| 9 | 85260 |
| 10 | 84914 |
| 11 | 84770 |
| 12 | 85651 |
| 13 | 86780 |
| 14 | 86008 |
| 15 | 91584 |

| Age | Number Still Alive | Remaining Life Expectancy |
|-----|--------------------|---------------------------|
| 0 | 100,000 | 77.7 |
| 1 | 99,329 | 77.2 |
| 2 | 99,285 | 76.3 |
| 3 | 99,255 | 75.3 |
| 4 | 99,233 | 74.3 |
| 5 | 99,216 | 73.3 |
| 6 | 99,199 | 72.3 |
| 7 | 99,184 | 71.3 |
| 8 | 99,169 | 70.4 |
| 9 | 99,157 | 69.4 |
| 10 | 99,147 | 68.4 |
| 11 | 99,138 | 67.4 |
| 12 | 99,130 | 66.4 |
| 13 | 99,117 | 65.4 |
| 14 | 99,097 | 64.4 |

Year of Cohort: 2007

¹³⁵ Arias, E. (2010, June). *United States life tables, 2006* (National Vital Statistics Reports vol. 58, no. 21). Washington, DC: United States Department of Health and Human Services, National Vital Statistics System, Table 1.

D10.4 Valuation of Reductions in Mortality Risk: Value of a Statistical Life

Several of the outcomes analyzed in the Institute's benefit-cost model affect the risk of mortality. For example, as described in Appendix D.4, if a prevention program reduces the risk that a participant will have a DSM alcohol disorder, then there is evidence that there will also be a reduced risk of an earlier-than-expected death.

The benefit-cost model employs two procedures to monetize the change in mortality risk.¹³⁶

The first procedure is sometimes called the "human capital" approach. This approach estimates the present value of lifetime labor market earnings that are lost because of an early death. In addition to lost labor market earnings, analysts sometimes include values of lost household production, valued at labor market rates, in the event of a death. As described in other sections of this Appendix, the Institute's model computes estimates for these lost human capital values using standard present-value procedures.

While the human capital approach places a monetary value of lost labor production, it does not provide an overall estimate of how much people would be willing to pay (or accept) for changes in mortality risk. To address this broader perspective, economists have been developing empirical estimates of the monetary value that people place on their lives. The general approach entails computing the value of a statistical life (VSL).¹³⁷ The VSL estimates are almost always much larger than the lost earnings from the human capital approach because VSL measures the total monetary value that people place on reduced risks of death, or the amounts that they are willing to accept for increased levels of mortality risk, and lost labor market earnings are only a portion of those valuations.

There are two general approaches used to calculate VSL: (1) the *revealed preferences* estimated from compensating wage differentials, and (2) the *stated preferences* elicited from people in surveys on how much they would be willing to pay to reduce the risk of death. Both approaches are active areas of current research and, among the more recent studies, the two approaches have been producing estimates that include quite similar ranges. Cropper, et al. (2011) reviewed both approaches and found that the revealed preference studies produce estimates of \$2.0 million to \$11.1 million (2009 USD), and that the stated preference studies produce VSL's in the range of \$2.0 million to \$8.0 million (2009 USD).

In addition to the current research on the calculation of an overall VSL, researchers are focusing on the heterogeneity of VSL by age and by risk level. Aldy and Viscusi (2008), after constructing revealed preference wage equations, have provided recent estimates of VSL for ages 18 to 62.¹³⁸ And Hammitt and Haninger (2010) have used a stated preference approach to estimate the VSL that adults place on children, compared to the VSL they state for adults.¹³⁹

The Institute's current approach to VSL includes specifying a range of VSLs to be used with Monte Carlo simulation, and applying the results from Aldy and Viscusi (2008) and Hammitt and Haninger (2010) to distribute VSL to individual years of a person's life. After computing these values, we then compute an adjusted VSL after subtracting the separately estimated "human capital" derived benefits of changes to lifetime earnings (LTE) and household production (HP). Thus, the general approach is:

$$VSL_{Adj} = VSL - LTE - HP$$

The Institute's VSL model is driven with the parameters shown in Exhibit D10.d, along with life table information displayed in Exhibit D10.c.

The user can specify a high, modal, and low value for VSL. These estimates are then modeled with a random draw from a triangular probability density distribution. For high and low VSL values, we use the preferred estimates reported in Kniesner et al (2011).¹⁴⁰ For the modal value, we compute the average between the high and low. These values are expressed in year 2001 dollars, and the model updates these values with the Implicit Price Deflator for Personal Consumption Expenditures to the user-selected base year for the benefit-cost model.

¹³⁶ For a general review of the analytical methods economists and others have used to assess the valuation of mortality risk, see Viscusi, W. I. (2008, March 18). *How to value a life* (Vanderbilt Law and Economics Research Paper No. 08-16), Nashville, TN: Vanderbilt University, Department of Economics.

¹³⁷ A recent review of the development of this research literature is provided in Cropper, M., Hammitt, J., & Robinson, L. (2011). *Valuing mortality risk reductions: Progress and challenges* (Working Paper No. 16971), Cambridge: National Bureau of Economic Research.

¹³⁸ Aldy, J. E., & Viscusi, W. K. (2008). Adjusting the value of a statistical life for age and cohort effects, *The Review of Economics and Statistics*, 90(3), 573-581.

¹³⁹ Hammitt, J. K., & Haninger, K. (2010). Valuing fatal risks to children and adults: Effects of disease, latency, and risk aversion, *Journal of Risk and Uncertainty*, 40(1), 57-83.

¹⁴⁰ Kniesner, T. J., Viscusi, W. K., & Ziliak, J. P. (2010). Policy relevant heterogeneity in the value of a statistical life: New evidence from panel data quantile regressions. *Journal of Risk and Uncertainty*, 40(1), 15-31.

The value of a statistical life year, VS_{LY}, is then computed for the range of years considered in the Kniesner study (ages 18 to 62) with the following equation, where the discount rate selected by the user is *disrate* and the average number of years of remaining life (for those currently 18 to 62) is taken from the general life table reported in Exhibit D10.c.

$$VS_{LY} = \frac{disrate \times VSL}{1 - (1 + disrate)^{-L}}$$

For example, with a \$7 million VSL (in 2001 dollars), a 3 percent discount rate, and 41 years of remaining life, the VS_{LY} is \$299,000 on average over the ages of 18 to 62. The next set of parameters in Exhibit D10.d are used to distribute this average VS_{LY} value over the different years of a person's life. We use the estimates from Aldy and Viscusi (2008) to compute a third-order polynomial (the parameters are shown on the user input sheet). The Aldy and Viscusi analysis, using revealed preference data from labor market wages, estimates the annual VS_{LY} for ages 18 to 62. Thus, by applying the third order polynomial to the base value (\$299,000) the following distributed estimates of VS_{LY} are obtained for ages 18 to 62.

Exhibit D10.d

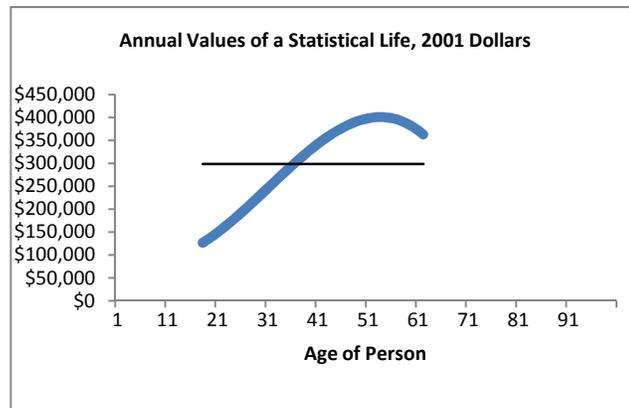
The screenshot shows the 'General' tab of the WSIPP Benefit-Cost Model. The interface includes a sidebar with various input categories and a main panel for parameter entry. The 'VSL' sub-tab is active, displaying parameters for estimating the value of a statistical life year and a table of public medical and social security costs.

Parameters to estimate the value of a statistical life year, ages 1 to 100

| | |
|-----------|---|
| 7.0 | Modal value of statistical life, millions |
| 10.0 | High value of statistical life, millions |
| 4.0 | Low value of statistical life, millions |
| 2001 | Year of dollars |
| 0.437446 | intercept |
| -0.031871 | x |
| 0.002142 | x^2 |
| -0.000023 | x^3 |
| -0.01 | post age 62 exponential change rate |
| 1.7 | Pre-age 18 multiplier |

Public Medical and Social Security Costs (Average cost per person)

| Age | Medical Costs | Social Security Payments |
|-----|---------------|--------------------------|
| 1 | 1091.809532 | 0 |
| 2 | 904.2224094 | 0 |
| 3 | 269.8754812 | 0 |
| 4 | 427.5071627 | 0 |
| 5 | 446.0323353 | 0 |
| 6 | 300.7446602 | 0 |
| 7 | 448.0675456 | 0 |
| 8 | 315.9048728 | 0 |
| 9 | 368.681332 | 0 |
| 10 | 365.0389391 | 0 |
| 11 | 233.7250909 | 0 |
| 12 | 239.8318219 | 0 |
| 13 | 297.0061829 | 0 |
| 14 | 205.3012262 | 0 |
| 15 | 416.9864218 | 0 |



The Aldy and Viscusi estimates only allow a distribution for ages 18 to 62. For ages older than 62, the empirical evidence is weak or non-existent. For these estimates, we follow the general approach taken by Viscusi and Hersch¹⁴¹ (2008) and apply values for older ages based on the values for the last years (around age 60 to 62) for which estimates are available. The parameter in Exhibit D10.d allows for an exponential rate of annual change that is multiplied by the age 62 value for VSLY. If zero is entered for the rate of change, then the VSLY value for age 62 is applied for all ages to 100. Thus, for ages 63 to 100, VSLY is computed with:

$$VSLY_y = VSLY_{62} \times (1 + esc)^{(y-62+1)}$$

For ages less than 18 (the earliest age for which a VSLY can be estimated with the Kniesner and Viscusi data), we use the ratio of VSL for children relative to adults reported in the stated preference paper by Hammitt and Haninger (2010). They found that the willingness to pay estimates for VSL for children are \$12 to \$15 million and \$6 to \$10 million for adults. We computed a point estimate for the ratio as $1.7 = (12 + 15)/2$ divided by $(6 + 10)/2$. In the model, this ratio is applied to the average adult VSLY. Thus, for ages 1 to 18, VSLY is computed with the Hammitt and Haninger ratio (HHratio):

$$VSLY_y = VSLY \times HHratio$$

¹⁴¹ Viscusi, W. K., & Hersch, J. (2008). The mortality cost to smokers. *Journal of Health Economics*, 27(4), 943-958.

D10.5 Deadweight Cost of Taxation

The model can compute estimates of the deadweight costs of taxation. The resulting values reflect the dollars of economic welfare loss per tax dollar raised to pay for program costs, or avoided if a program reduces taxpayer financed costs.¹⁴² Because there is uncertainty around the appropriate values of deadweight costs, we model low, modal, and high multiplicative values. When the model is run in non-simulation mode, the modal deadweight value is used. In Monte Carlo simulation, each run randomly draws a deadweight value from a triangular probability density distribution, with the user-selected low, modal, and high deadweight values defining the triangle. The deadweight cost value is then multiplied by any tax-related cost or tax-related benefit of the program. The resulting net deadweight cost values are tallied and reported in the “Other Benefits” section of the output. For example, if a program costs taxpayers \$1,000 per participant, and it is estimated that the program saves \$600 in taxpayer savings from an improved outcome, e.g., less taxpayer spending on the criminal justice system, then with a modal deadweight cost value of 50 percent, there would be a net deadweight cost of the program of \$200 (\$600 times 50% minus \$1,000 times 50%). In the actual run of the model, these calculations are carried out for each year of cash flows.

$$(D12.1) \text{ } DWL_{age} = \sum_{y=age}^N \frac{(B_y - C_y) \times DWL\%}{(1 + Dis)^y}$$

Exhibit D10.e is a screen shot showing where the three deadweight cost values are entered. The Institute uses a low real deadweight cost value of 0 percent, a modal rate of 50 percent, and a high rate of 100 percent. These input choices are the same values used by Heckman et al. (2010) in their analysis of the benefits and costs of the Perry Preschool program.¹⁴³

Exhibit D10.e

The screenshot shows the 'WSIPP Benefit-Cost Model: Version 1.1' window. The 'Enter Program Inputs' tab is active, and the 'Deadweight Cost' sub-tab is selected. The main area displays the text 'Deadweight Cost is dollar of welfare loss per tax dollar.' Below this, three input fields are shown: 'Low' with a value of 0, 'Medium' with a value of 0.5, and 'High' with a value of 1. A 'Back to Main Model' button is located at the top left of the input area. The left sidebar contains various category buttons such as 'General', 'Economic', 'Crime', 'Education', 'Child Welfare', 'Substance Use', 'Health Care', 'Mental Health', 'Public Asst', 'Housing', 'Teen Birth', and 'Outcomes & Links'.

¹⁴² Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (1996). *Cost-benefit analysis: Concepts and practice* (4th ed). Upper Saddle River, NJ: Prentice Hall.

¹⁴³ Heckman et al., 2010

D10.6 Inflation/Price Indexes

As noted, many of the monetary values in the model are denominated in different years' monetary units. The model converts each of these to the base year chosen by the user. Exhibit D10.f displays the input screen where the price indices used by the model are entered. The general inflation index used by the Institute is United States Department of Commerce's Chain-Weighted Implicit Price Deflator for Personal Consumption Expenditures. The forecast years for the index is taken from the Washington State Economic and Revenue Forecast Council, the official forecasting agency for Washington State government. Since health care costs are central in the Institute's benefit-cost model, and since health care prices have followed different paths than general prices, we also include a medical cost index, as shown in the Exhibit. We use the Medical Care Index of the Consumer Price Index for all urban consumers, published by the United State Department of Labor.

Exhibit D10.f

The screenshot shows the 'Economic' input screen of the WSIPP Benefit-Cost Model. The interface includes a navigation pane on the left with buttons for 'General', 'Economic', 'Crime', 'Education', 'Child Welfare', 'Substance Use', 'Health Care', 'Mental Health', 'Public Asst', 'Housing', 'Teen Birth', and 'Outcomes & Links'. The 'Economic' section is active, showing a 'Back to Main Model' button and tabs for 'Inflation Index', 'Earnings & Benefits', 'Misc.', and 'HouseHold Production'. Two data tables are displayed:

| Implicit Price Deflator for Personal Consumption Expenditures | |
|---|-------------|
| Year | Index Value |
| 1960 | 0.186 |
| 1961 | 0.188 |
| 1962 | 0.190 |
| 1963 | 0.192 |
| 1964 | 0.195 |
| 1965 | 0.198 |
| 1966 | 0.203 |
| 1967 | 0.208 |
| 1968 | 0.216 |
| 1969 | 0.226 |
| 1970 | 0.237 |
| 1971 | 0.247 |
| 1972 | 0.255 |
| 1973 | 0.269 |
| 1974 | 0.297 |

| CPI All Urban Consumers, Medical Care | |
|---------------------------------------|-------------|
| Year | Index Value |
| 1983 | |
| 1984 | 106.800 |
| 1985 | 113.500 |
| 1986 | 122.000 |
| 1987 | 130.100 |
| 1988 | 138.600 |
| 1989 | 149.300 |
| 1990 | 162.800 |
| 1991 | 177.000 |
| 1992 | 190.100 |
| 1993 | 201.400 |
| 1994 | 211.000 |
| 1995 | 220.500 |
| 1996 | 228.200 |
| 1997 | 234.600 |

D10.7 Household Production

In addition to the value of reduced or lost labor market value in the commercial economy, many studies of morbidity and mortality costs include estimates of the reduced or lost value of household production. We adopt that approach in this study. The model computes the value of lost household production that might be shifted to another in the event of death. Monetizing the value of household production is a common procedure in cost-of-illness studies.¹⁴⁴ We estimate 19.5 hours per week for household production. This estimate is based on an assumed 1.5 hours per day for housekeeping services, 1.0 hours per day for food preparation, and 2.0 hours per week for household maintenance. These estimates are quite close to the 21.4 hours per week calculated by Douglass et al.¹⁴⁵ The average shadow wage rate for these three household services was taken from United State Bureau of Labor Statistics data on average wage rates in Washington in 2004 for each service¹⁴⁶

Exhibit D10.g

| | |
|---------|---|
| 19.5 | Hours per week |
| 10.08 | Dollars per hour |
| 2004 | year of dollars |
| 0.4273 | Shift parameter intercept |
| 0.01831 | Shift parameter x |
| -0.0002 | Shift parameter x^2 |
| 18 | Year to begin the shift process |
| 0.1 | Annual probability that a someone re-attaches to someone else following death of spouse |

To compute the household production effect for the incidence of the DSM disorders, we begin with the following equation:

$$H_a = HOURS * \$HOURL * 52 * PrSHIFT_a * INFLATION$$

¹⁴⁴ See, for example, Max, W., Rice, D., Sung, H., & Michel, M. (2004). *Valuing human life: Estimating the present value of lifetime earnings, 2000* (Paper PVLE2000). San Francisco: University of California, San Francisco. Retrieved June 30, 2011 from <http://escholarship.org/uc/item/82d0550k#page-1>

¹⁴⁵ Douglass, J., Kenney, G., & Miller, T. (1990). Which estimates of household production are best? *Journal of Forensic Economics*, 4(1), 25-45.

¹⁴⁶ Bureau of Labor Statistics. *November 2004 Occupational employment and wage estimates*. Retrieved June 30, 2011 from http://www.bls.gov/oes/current/oes_wa.htm#b39-0000

Not all of the value of lost household production will be shifted to others if a person dies or is disabled as a result of having an alcohol, drug, or mental health disorder. Some people live alone and no one else is required to assume the household production if the person becomes disabled or dies as a result of the disorder. We provide an estimate for this with the variable $PrSHIFT_a$, used in the previous equation. This variable provides an estimate of the probability that a person at age (a) will not be living alone and, if he or she becomes disordered, that the value of his or her household production will be shifted to someone else. We estimate this probability with national data from the same Current Population Survey described above. The results of this estimation and are computed with this equation:

$$Pr SHIFT_a = \frac{FHH_a}{(HH_a - GQ_a)}$$

The probability of shifting household production $PrSHIFT_a$ in the event of a disorder is given by the total number of people in households with family members (FHH_a) divided by the total number of people in households (HH_a) (less those living in group quarters (GQ_a)). Values for all three variables come from the CPS.

The annual cash flows of lost household production associated with having a disorder of type t is estimated with this process:

$$\$HP_{ty} = \sum_p^P H_{p+y-1} * (1 + ER)^{y-1} * EE_t * PP_{tp} * -1$$

In this equation, $\$HP_{ty}$ is the annual cash flow of shifted household production in year y , where y is the number of years following participation in a program.

D10.8 Tax Rates

The benefit cost model uses average tax rates¹⁴⁷ for several calculations. We use the aggregate total from the Tax Foundation to represent a combination of all kinds (income, sales, property, and other) of taxes paid, as a percentage of income. This value is entered on the screen shot displayed in Exhibit10.h.

D10.9 Capital Costs

A few routines in the model use capital financing costs. The real cost of capital was obtained from discussions with fiscal staff of the Washington State legislature. This value is entered on the screen shot displayed in Exhibit10.h.

Exhibit D10.h

WSIPP Benefit-Cost Model: Version 1.1

Enter Sector Inputs Enter Program Inputs Run Models & View Reports

General Economic Crime Education Child Welfare Substance Use Health Care Mental Health Public Asst Housing Teen Birth Outcomes & Links

Economic

Back to Main Model

Inflation Index Earnings & Benefits Misc. HouseHold Production

Real Cost of Capital

0.05

Total Effective Tax Rate

0.269

¹⁴⁷ Padgitt, K. M., (2011, March). *Tax Freedom Day@ arrives on April 12* (Table 1, p. 3) (Special Report No. 190). Washington, DC: Tax Foundation. Retrieved June 30, 2011 from: <http://www.taxfoundation.org/files/sr190.pdf>.

Appendix E: Meta Analyses of Linked Outcomes

E1. Input Screen for Linked Outcome Effect Sizes

One of the features of the Institute’s benefit-cost model is its use of empirically-established causal “links” between two outcomes. The logic follows this path: If a program evaluation establishes a causal effect of program *P* on outcome *O1*, and another body of research measures a causal relationship between outcome *O1* and outcome *O2*, then it logically follows that *P* must have an effect on *O2*.

$$\text{if } P \rightarrow O1, \quad \text{and } O1 \rightarrow O2, \quad \text{then } P \rightarrow O2$$

For example, if the juvenile justice program Functional Family Therapy (FFT) is shown to affect juvenile crime outcomes, and if separately analyzed longitudinal research establishes that juvenile crime is causally related to high school graduation probability, then FFT can be assumed to have an effect on high school graduation. Thus, while none of the outcome evaluations of in our meta-analytic review of FFT (see Appendix I.) measured the effect of the program on high school graduation, it is reasonable to assume that there is a relationship between FFT and high school graduation since there is a separate body of research that demonstrates the linkage between juvenile crime and high school graduation.

The purpose of the Institute’s analyses of linked outcomes is to take advantage of this additional information. This is especially important in conducting benefit-cost analysis where the focus is on long-term effects from (usually) short-term program evaluations. Exhibit E1 displays a screen shot of where the linked effect size and standard error are entered. Exhibit E2 displays the Institute’s current meta analytic results of the linkage literature, and Exhibit E3 lists the individual studies that were meta-analyzed to establish the causal estimates.

Exhibit E1

The screenshot shows the 'Outcomes and Links' section of the WSIPP Benefit-Cost Model. The main window contains a table of outcomes and a detailed view of the selected 'Crime' outcome.

| WSIPP Outcome Number | Outcome Name (these are all the program outcomes that our model can do something with) | Outcome Display Location | Dichot. or Continuous Outcome |
|----------------------|--|--------------------------|-------------------------------|
| 4 | Child abuse and neglect | 1 | D |
| 10 | Out-of-home placement | 2 | D |
| 1 | Crime | 3 | D |
| 2 | High school graduation | 4 | D |
| 3 | Test scores | 5 | C |
| 7 | K-12 special education | 6 | D |
| 6 | K-12 grade repetition | 7 | D |
| 8 | Years of education | 8 | C |
| 9 | Age of initiation (tobacco) | 9 | C |
| 12 | Regular smoking | 10 | D |
| 13 | Age of initiation (alcohol) | 11 | C |
| 15 | Alcohol abuse or dependence | 12 | D |
| 11 | Age of initiation (cannabis) | 13 | C |
| 1 | Crime | 3 | D |

| The effect of this selected outcome, ... on the following monetization area: | ES of Outcome on Money | SE of ES of Outcome on Money | Age at which relationship begins | Age at which relationship is measured |
|--|------------------------|------------------------------|----------------------------------|---------------------------------------|
| Crime | 1 | 1 | 1 | 1 |
| K-12 system: year of education | | | | |
| K-12 system: special education | | | | |
| K-12 system: grade repetition | | | | |
| Child abuse and neglect | | | | |
| Earnings via high school graduation | -0.341 | 0.084 | 18 | 18 |
| Earnings via test scores | | | | |
| Earnings: Years in school | | | | |
| Earnings: | | | | |
| Earnings: Crime | | | | |
| Earnings: Tobacco, Regular Use | | | | |
| Out-of-home placement | | | | |
| Earnings: Morbidity | | | | |
| Earnings: DSM Alcohol Disorder | | | | |
| Property Loss: Alcohol | | | | |
| Health Care Costs: Alcohol | | | | |
| Health Care Costs: Tobacco | | | | |
| Earnings: DSM Cannabis Disorder | | | | |
| Earnings: DSM Illicit Drug Disorder | | | | |
| Health Care Costs: Illicit Drugs | | | | |
| Property Loss: Illicit Drugs | | | | |
| Health Care Costs: Cannabis | | | | |
| Earnings: DSM Depression | | | | |

E2. Institute Adjustments to Effect Sizes for Methodological Quality, Generalizability of the Sample, and Relevance of the Independent and Dependent Variables

In the last two columns of Exhibit E2 we list the “Adjusted Effect Size” and standard error that we use in our analyses. As we do with the results from program evaluations, we make adjustments to the initial effect sizes to account for various forms of unobserved heterogeneity that we suspect exists in the underlying studies. We make three types of adjustments that we deem to be necessary to increase our confidence in the evidence for a causal relationship between two outcomes. We make adjustments for: (a) the methodological quality of each study we include in the meta-analyses; (b) the degree to which findings for a particular sample of people can be generalized to other populations; and (c) the relevance of the independent and dependent measures that individual studies examined.

E2.1 Methodological Quality

As we do with the program evaluation literature, we also apply weights to studies to account for expected biases that probably exist in certain types of research designs. To establish that one outcome leads to another, longitudinal studies that establish temporal ordering—that a first outcome (e.g., juvenile crime) precedes another outcome (e.g., high school graduation)—and include measures of other factors that also influence the outcome of interest are preferred. Ideally, the study would statistically control for both observable factors—e.g., family income—and unobservable variables by using fixed effects modeling, natural experiments, twin studies, instrumental variables, or other techniques. Other studies may be cross-sectional or may not statistically control for as many other potentially confounding factors; this does not mean that results from these studies are of no value, but it does mean that less confidence can be placed in any cause-and-effect conclusions drawn from the results.

To account for the differences in the quality of research designs, we use a 6-point scale (with values ranging from zero to five) as a way to adjust the reported results. On this scale, a rating of “95” reflects a study in which the most confidence can be placed: a longitudinal study with temporal ordering and good controls for observable and unobservable confounds. A rating of “90” reflects a study in which temporal ordering is not established and we cannot infer a causal link between independent and dependent variables.

On the 90-to-95 scale as interpreted by the Institute, each study is rated as follows.

- A “95” is assigned to a longitudinal study with temporal ordering and good statistical controls for observable AND unobservable confounds.
- A “94” is assigned to a longitudinal study with temporal ordering and good statistical controls for observable confounds.
- A “93” is assigned to a longitudinal study with temporal ordering and not as many controls.
- A “92” is assigned to a cross-sectional study with temporal ordering, and retrospective measurement.
- A “91” is a placeholder rating that is not currently used.
- A “90” involves a study in which we cannot infer causal link between independent and dependent variables.

In our meta analyses, we do not use the results from studies rated as a “90” or “91” on this scale.

An explicit adjustment factor is assigned to the results of individual effect sizes based on the Institute’s judgment concerning research design quality. This adjustment is critical and the only practical way to combine the results from high quality studies (e.g., a level 95 study) with those of lesser design quality (level 94 and lower studies).

The effect of the adjustment is to multiply the effect size for any study by the appropriate research design factor. For example, if a study has an effect size of -.20 and it is deemed a level 4 study, then the -.20 effect size would be multiplied by .75 to produce a -.15 adjusted effect size for use in the analysis. In Exhibit E3, the column labeled “research design” indicates the discount applied to each study’s results used in the meta-analyses.

E2.2 Generalizability of the Sample

We also adjust the effect sizes for linked outcomes for the degree to which the individuals included in the study sample are representative of the population as a whole. If we determine that a sample is not representative of the Washington state population, we use a multiplicative factor of .75 to adjust the effect size downward.

E2.3 Relevance of the Independent and Dependent Variables

Some studies use outcome measures that may not be precise gauges of the way the benefit-cost model monetizes results. In these cases, we record a flag that can later be used to discount the effect. For example, the benefit-cost model monetizes disordered alcohol use based on a DSM-level alcohol disorder. If a longitudinal study measures a linkage between “heavy drinking” (but not DSM alcohol use) and employment, then we will flag this weaker measure. In these cases, we discount the effect sizes by set factors.

Exhibit E2
Linked Outcomes
Meta-Analytic Estimates of Standardized Mean Difference Effect Sizes

| Estimated Causal Links Between Outcomes | Topic Reference Number | Number of Effect Sizes | Meta-Analytic Results Before Adjusting Effect Sizes | | | | | | | | | Adjusted Effect Size and Standard Error Used in the Benefit-Cost Analysis | |
|--|------------------------|------------------------|---|---------|-------|--|---------|-------------------------------------|-------|-------|-------|---|--|
| | | | Fixed Effects Model | | | | | Random Effects | | | | | |
| | | | Weighted Mean Effect Size & p-value | | | Homogeneity Test (p-value to reject homogeneity) | | Weighted Mean Effect Size & p-value | | | ES | SE | |
| ES | SE | p-value | Q-stat | p-value | ES | SE | p-value | ES | SE | | | | |
| Child Abuse & Neglect, leading to... | | | | | | | | | | | | | |
| High school graduation | 1 | 5 | -.412 | 0.048 | 0.000 | 14.31 | 0.006 | -.404 | 0.098 | 0.000 | -.212 | 0.098 | |
| Special education | 5 | 1 | .389 | 0.036 | 0.000 | 0.00 | na | | | | .194 | 0.036 | |
| Test scores-academic | 2 | 3 | -.248 | 0.052 | 0.000 | 0.86 | 0.649 | -.248 | 0.052 | 0.000 | -.099 | 0.052 | |
| Alcohol (disordered use) | 3 | 5 | .163 | 0.028 | 0.000 | 3.08 | 0.545 | .163 | 0.028 | 0.000 | .055 | 0.028 | |
| Illicit drugs (disordered use) | 4 | 5 | .279 | 0.061 | 0.000 | 2.21 | 0.697 | .279 | 0.061 | 0.000 | .200 | 0.061 | |
| Depression | 6 | 7 | .289 | 0.029 | 0.000 | 16.81 | 0.010 | .257 | 0.053 | 0.000 | .095 | 0.053 | |
| Teen births <18 | 7 | 2 | .200 | 0.113 | 0.076 | 7.26 | 0.007 | .431 | 0.403 | 0.285 | .175 | 0.403 | |
| Any crime measure | 8 | 12 | .503 | 0.027 | 0.000 | 56.45 | 0.000 | .513 | 0.069 | 0.000 | .254 | 0.069 | |
| Alcohol Disorder, leading to... | | | | | | | | | | | | | |
| Any crime measure | 11 | 6 | .237 | 0.017 | 0.000 | 204.26 | 0.000 | .262 | 0.124 | 0.035 | .134 | 0.124 | |
| Employment | 12 | 13 | -.514 | 0.012 | 0.000 | 1228.76 | 0.000 | -.464 | 0.128 | 0.000 | -.210 | 0.128 | |
| Earnings/Wages | 13 | 9 | -.082 | 0.012 | 0.000 | 30.45 | 0.000 | -.065 | 0.025 | 0.009 | -.048 | 0.025 | |
| Illicit Drug Disorder, leading to... | | | | | | | | | | | | | |
| Employment | 14 | 5 | -.211 | 0.026 | 0.000 | 3.18 | 0.528 | -.211 | 0.026 | 0.000 | -.126 | 0.026 | |
| Earnings/Wages | 15 | 2 | -.048 | 0.033 | 0.149 | 27.82 | 0.000 | -.333 | 0.343 | 0.333 | -.166 | 0.343 | |
| Any crime measure | 16 | 1 | .250 | 0.100 | 0.013 | 0.00 | na | | | | .187 | 0.100 | |
| Cannabis use, leading to... | | | | | | | | | | | | | |
| High school graduation | 17 | 13 | -.205 | 0.012 | 0.000 | 71.48 | 0.000 | -.286 | 0.038 | 0.000 | -.211 | 0.038 | |
| Smoking, leading to... | | | | | | | | | | | | | |
| Employment | 18 | 2 | -.037 | 0.009 | 0.000 | 19.16 | 0.000 | -.083 | 0.062 | 0.177 | -.062 | 0.062 | |
| Earnings/Wages | 19 | 5 | -.040 | 0.006 | 0.000 | 247.68 | 0.000 | -.128 | 0.067 | 0.056 | -.067 | 0.067 | |
| Mental Health Disorder, leading to... | | | | | | | | | | | | | |
| Employment (DSM MI) | 27 | 18 | -.409 | 0.007 | 0.000 | 883.94 | 0.000 | -.406 | 0.059 | 0.000 | -.197 | 0.059 | |
| Earnings/Wages (DSM MI) | 28 | 9 | -.081 | 0.007 | 0.000 | 35.73 | 0.000 | -.095 | 0.020 | 0.000 | -.060 | 0.020 | |
| Employment (Depression) | 29 | 10 | -.314 | 0.018 | 0.000 | 80.59 | 0.000 | -.349 | 0.060 | 0.000 | -.174 | 0.060 | |
| Earnings/Wages (Depression) | 30 | 4 | -.043 | 0.020 | 0.028 | 1.95 | 0.582 | -.043 | 0.020 | 0.028 | -.030 | 0.020 | |
| Employment (Anxiety disorder) | 31 | 7 | -.141 | 0.019 | 0.000 | 57.47 | 0.000 | -.231 | 0.068 | 0.001 | -.110 | 0.068 | |
| Earnings/Wages (Anxiety disorder) | 32 | 3 | -.205 | 0.027 | 0.000 | 35.14 | 0.000 | -.187 | 0.116 | 0.108 | -.126 | 0.116 | |
| Any crime measure (Externalizing composite) | 33 | 12 | .316 | 0.027 | 0.000 | 26.24 | 0.006 | .334 | 0.048 | 0.000 | .230 | 0.048 | |
| High school graduation (Externalizing composite) | 34 | 17 | -.323 | 0.015 | 0.000 | 36.65 | 0.002 | -.343 | 0.028 | 0.000 | -.229 | 0.028 | |
| High school graduation (Internalizing composite) | 35 | 11 | -.091 | 0.016 | 0.000 | 24.25 | 0.007 | -.113 | 0.034 | 0.001 | -.066 | 0.034 | |
| Grade retention (ADHD) | 36 | 4 | .466 | 0.048 | 0.000 | 3.02 | 0.388 | .466 | 0.049 | 0.000 | .321 | 0.049 | |
| Test scores-academic (ADHD) | 37 | 3 | -.370 | 0.039 | 0.000 | 14.09 | 0.001 | -.375 | 0.102 | 0.000 | -.282 | 0.102 | |
| High school graduation (ADHD) | 38 | 6 | -.301 | 0.023 | 0.000 | 1.90 | 0.863 | -.301 | 0.023 | 0.000 | -.213 | 0.023 | |
| Any crime measure (ADHD) | 39 | 5 | .273 | 0.039 | 0.000 | 17.37 | 0.002 | .316 | 0.096 | 0.001 | .168 | 0.096 | |
| High school graduation (Conduct Disorder) | 40 | 7 | -.426 | 0.028 | 0.000 | 9.58 | 0.144 | -.443 | 0.045 | 0.000 | -.277 | 0.045 | |
| Any crime measure (Conduct Disorder) | 41 | 7 | .355 | 0.037 | 0.000 | 6.54 | 0.365 | .356 | 0.041 | 0.000 | .275 | 0.041 | |
| Any crime measure (Schiz. or Bipolar, Inpatient) | 42 | 3 | .525 | 0.029 | 0.000 | 114.70 | 0.000 | .529 | 0.263 | 0.044 | .265 | 0.263 | |

Exhibit E2, continued

| Estimated Causal Links Between Outcomes | Topic Reference Number | Number of Effect Sizes | Meta-Analytic Results Before Adjusting Effect Sizes | | | | | | | | | Adjusted Effect Size and Standard Error Used in the Benefit-Cost Analysis | |
|--|------------------------|------------------------|---|-------|---------|--|---------|-------------------------------------|-------|---------|----|---|--------------|
| | | | Fixed Effects Model | | | | | Random Effects | | | | | |
| | | | Weighted Mean Effect Size & p-value | | | Homogeneity Test (p-value to reject homogeneity) | | Weighted Mean Effect Size & p-value | | | ES | SE | |
| | | | ES | SE | p-value | Q-stat | p-value | ES | SE | p-value | | | |
| Teen Birth (< 18 years old), leading to... | | | | | | | | | | | | | |
| Child abuse and neglect (Births < 18, child) | 20 | 1 | .238 | 0.008 | 0.000 | 0.00 | na | | | | | .119 | 0.008 |
| Out-of-home placements (Births < 18, child) | 21 | 1 | .116 | 0.013 | 0.000 | 0.00 | na | | | | | .058 | 0.013 |
| Grade retention (Births < 18, child) | 23 | 4 | .205 | 0.033 | 0.000 | 2.03 | 0.567 | .205 | 0.033 | 0.000 | | .202 | 0.033 |
| Any crime measure (Births < 18, child) | 26 | 2 | .183 | 0.068 | 0.007 | 0.70 | 0.403 | .183 | 0.068 | 0.007 | | .137 | 0.068 |
| High school graduation (Births < 18, child) | 24 | 3 | -.213 | 0.068 | 0.002 | 0.84 | 0.657 | -.213 | 0.068 | 0.002 | | -.127 | 0.068 |
| Public Assistance (Births < 18, mother) | 22 | 3 | .173 | 0.096 | 0.072 | 4.65 | 0.098 | .244 | 0.164 | 0.138 | | .145 | 0.164 |
| High school graduation (Births < 18, mother) | 25 | 4 | -.181 | 0.072 | 0.013 | 0.79 | 0.852 | -.181 | 0.072 | 0.013 | | -.155 | 0.072 |
| High School Graduation, leading to... | | | | | | | | | | | | | |
| Any crime measure | 10 | 11 | -.150 | 0.008 | 0.000 | 100.06 | 0.000 | -.215 | 0.030 | 0.000 | | -.143 | 0.030 |
| Crime, leading to... | | | | | | | | | | | | | |
| High school graduation | 9 | 6 | -.469 | 0.032 | 0.000 | 33.91 | 0.000 | -.496 | 0.091 | 0.000 | | -.393 | 0.091 |

Exhibit E3

| Individual Effect Sizes Used in the Meta-Analysis of Each Topic | | | | | | | | | | | |
|--|--------------|-------------------------|--------------------------------|-------------------------|---------------------------------------|--|---|-----------------------------|--|--|----------------------|
| Study (Date) | Topic Number | Study Results | | | | | Multiplicative Weights & Adjusted Effect Size | | | | |
| | | Un-adjusted effect size | Number in test condition group | Number in control group | Inverse variance weight-fixed effects | Inverse variance weight-random effects | Re-search design | General-izability of sample | Relevance of inde-pendent variable measure | Relevance of de-pendent variable measure | Adjusted effect size |
| Thornberry et al., 2001 | 1 | -0.176 | 134 | 604 | 45.9 | 18.2 | 0.75 | 1.00 | 1.00 | 1.00 | -0.132 |
| McGloin & Widom, 2001 | 1 | -0.479 | 676 | 520 | 185.5 | 25.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.240 |
| Lansford et al., 2007 | 1 | -0.854 | 69 | 505 | 34.5 | 16.0 | 0.75 | 1.00 | 0.50 | 1.00 | -0.320 |
| Boden et al., 2007 | 1 | -0.158 | 171 | 800 | 64.5 | 20.5 | 0.75 | 1.00 | 0.50 | 1.00 | -0.059 |
| Mersky & Topitzes, 2010 | 1 | -0.407 | 179 | 1148 | 99.5 | 23.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.305 |
| Lansford et al., 2002 | 2 | -0.145 | 50 | 387 | 44.2 | 44.2 | 0.75 | 1.00 | 0.50 | 1.00 | -0.054 |
| Slade & Wissow, 2007 | 2 | -0.286 | 632 | 1146 | 209.6 | 209.6 | 1.00 | 1.00 | 0.50 | 0.50 | -0.071 |
| Topitzes et al., 2010 | 2 | -0.220 | 135 | 990 | 118.5 | 118.5 | 0.75 | 1.00 | 1.00 | 1.00 | -0.165 |
| Fergusson & Lynskey, 1997 | 3 | 0.409 | 118 | 111 | 23.9 | 23.9 | 0.50 | 1.00 | 0.50 | 1.00 | 0.102 |
| Scott et al., 2010 | 3 | 0.200 | 221 | 1923 | 47.6 | 47.6 | 0.50 | 1.00 | 1.00 | 1.00 | 0.100 |
| Thornberry et al., 2010 | 3 | 0.171 | 170 | 645 | 134.2 | 134.2 | 0.75 | 1.00 | 1.00 | 1.00 | 0.129 |
| Horwitz et al., 2001 | 3 | 0.075 | 637 | 510 | 192.3 | 192.3 | 0.50 | 1.00 | 1.00 | 1.00 | 0.038 |
| Shin et al., 2009 | 3 | 0.173 | 6729 | 6019 | 851.9 | 851.9 | 0.50 | 1.00 | 0.50 | 1.00 | 0.043 |
| Fergusson & Lynskey, 1997 | 4 | 0.176 | 118 | 111 | 15.7 | 15.7 | 0.50 | 1.00 | 0.50 | 1.00 | 0.044 |
| Scott et al., 2010 | 4 | 0.417 | 221 | 1923 | 31.5 | 31.5 | 0.50 | 1.00 | 1.00 | 1.00 | 0.208 |
| Thornberry et al., 2010 | 4 | 0.275 | 170 | 645 | 133.7 | 133.7 | 0.75 | 1.00 | 1.00 | 1.00 | 0.206 |
| Fergusson et al., 2008 | 4 | 0.113 | 162 | 839 | 38.9 | 38.9 | 0.50 | 1.00 | 0.50 | 1.00 | 0.028 |
| Arteaga et al., 2010 | 4 | 0.367 | 117 | 1091 | 48.6 | 48.6 | 1.00 | 1.00 | 1.00 | 1.00 | 0.367 |
| Jonson-Reid et al., 2004 | 5 | 0.389 | 3987 | 3953 | 767.9 | 767.9 | 0.50 | 1.00 | 1.00 | 1.00 | 0.194 |
| Chapman et al., 2004 | 6 | 0.411 | 2373 | 7087 | 511.7 | 72.4 | 0.50 | 1.00 | 0.50 | 1.00 | 0.103 |
| Widom et al., 2007 | 6 | 0.145 | 676 | 520 | 139.3 | 52.5 | 0.50 | 1.00 | 1.00 | 1.00 | 0.072 |
| Scott et al., 2010 | 6 | 0.366 | 221 | 1923 | 59.2 | 34.8 | 0.50 | 1.00 | 1.00 | 1.00 | 0.183 |
| Thornberry et al., 2010 | 6 | 0.158 | 170 | 645 | 134.3 | 51.8 | 0.75 | 1.00 | 1.00 | 1.00 | 0.118 |
| Fletcher, 2009 | 6 | 0.297 | 182 | 3840 | 80.4 | 41.2 | 1.00 | 1.00 | 0.50 | 1.00 | 0.148 |
| Springer et al., 2007 | 6 | 0.156 | 234 | 1817 | 207.0 | 59.9 | 0.50 | 1.00 | 0.50 | 0.50 | 0.020 |
| Fergusson et al., 2008 | 6 | 0.266 | 162 | 839 | 76.5 | 40.1 | 0.50 | 1.00 | 0.50 | 1.00 | 0.067 |
| Widom & Kuhns, 1996 | 7 | 0.063 | 338 | 244 | 65.4 | 3.4 | 0.75 | 1.00 | 1.00 | 1.00 | 0.047 |
| Noll et al., 2003 | 7 | 0.873 | 77 | 89 | 13.3 | 2.8 | 0.50 | 0.75 | 1.00 | 1.00 | 0.327 |
| Cohen et al., 2004 | 8 | 0.733 | 51 | 611 | 30.5 | 14.0 | 0.50 | 1.00 | 0.50 | 1.00 | 0.183 |
| Maxfield & Widom, 1996 | 8 | 0.272 | 908 | 667 | 253.0 | 23.4 | 0.50 | 1.00 | 1.00 | 1.00 | 0.136 |
| English et al., 2002 | 8 | 0.600 | 877 | 877 | 235.6 | 23.2 | 0.50 | 1.00 | 1.00 | 0.50 | 0.150 |
| Stouthamer-Loeber et al., 2001 | 8 | 0.379 | 52 | 104 | 23.1 | 12.2 | 0.50 | 1.00 | 1.00 | 1.00 | 0.190 |
| Fergusson & Lynskey, 1997 | 8 | 0.340 | 118 | 111 | 13.7 | 8.9 | 0.50 | 1.00 | 0.50 | 1.00 | 0.085 |
| Currie & Tekin, 2006 | 8 | 0.494 | 3121 | 10388 | 414.1 | 24.2 | 1.00 | 1.00 | 0.50 | 1.00 | 0.247 |
| Lansford et al., 2007 | 8 | 0.084 | 69 | 505 | 37.7 | 15.3 | 0.75 | 1.00 | 0.50 | 1.00 | 0.031 |
| Mersky & Reynolds, 2007 | 8 | 0.451 | 129 | 1275 | 49.3 | 16.9 | 0.75 | 1.00 | 1.00 | 1.00 | 0.338 |
| Lemmon, 1999 | 8 | 1.083 | 267 | 365 | 79.8 | 19.5 | 0.50 | 0.75 | 1.00 | 1.00 | 0.406 |
| Stouthamer-Loeber et al., 2002 | 8 | 0.635 | 83 | 140 | 28.3 | 13.5 | 0.75 | 1.00 | 1.00 | 1.00 | 0.476 |
| Thornberry et al., 2010 | 8 | 0.342 | 170 | 645 | 82.6 | 19.6 | 0.75 | 1.00 | 1.00 | 1.00 | 0.257 |
| Van Dorn, et al., 2011 | 8 | 0.620 | 3465 | 31188 | 161.4 | 22.2 | 0.75 | 1.00 | 1.00 | 1.00 | 0.465 |
| Hjalmarsson, 2008 | 9 | -0.183 | 467 | 6950 | 263.1 | 23.9 | 1.00 | 1.00 | 1.00 | 1.00 | -0.183 |
| Tanner et al., 1999 | 9 | -0.403 | 478 | 1882 | 130.9 | 21.9 | 0.75 | 1.00 | 1.00 | 1.00 | -0.302 |
| Hirschfield, 2009 | 9 | -0.666 | 216 | 2039 | 31.6 | 14.4 | 0.75 | 1.00 | 1.00 | 1.00 | -0.500 |
| Apel & Sweeten, 2009 | 9 | -0.621 | 400 | 4649 | 233.4 | 23.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.466 |
| Webbink et al., 2008 | 9 | -0.588 | 1126 | 1126 | 318.1 | 24.3 | 1.00 | 1.00 | 1.00 | 1.00 | -0.588 |
| Kirk & Sampson, 2009 | 9 | -0.637 | 76 | 102 | 26.4 | 13.2 | 0.50 | 1.00 | 1.00 | 1.00 | -0.319 |

Exhibit E3

| Individual Effect Sizes Used in the Meta-Analysis of Each Topic | | | | | | | | | | | |
|--|--------------|-------------------------|--------------------------------|-------------------------|---------------------------------------|--|---|-----------------------------|--|--|----------------------|
| Study (Date) | Topic Number | Study Results | | | | | Multiplicative Weights & Adjusted Effect Size | | | | |
| | | Un-adjusted effect size | Number in test condition group | Number in control group | Inverse variance weight-fixed effects | Inverse variance weight-random effects | Re-search design | General-izability of sample | Relevance of inde-pendent variable measure | Relevance of de-pendent variable measure | Adjusted effect size |
| Lochner & Moretti, 2004 | 10 | -0.083 | 16319 | 16319 | 8152.7 | 127.8 | 1.00 | 1.00 | 1.00 | 1.00 | -0.083 |
| Lochner & Moretti, 2004 | 10 | -0.458 | 1724 | 978 | 287.7 | 89.5 | 0.50 | 1.00 | 1.00 | 1.00 | -0.229 |
| Levitt & Lochner, 2001 | 10 | -0.144 | 2135 | 2153 | 706.9 | 109.7 | 0.50 | 1.00 | 1.00 | 1.00 | -0.072 |
| Sabates, 2008 | 10 | -0.200 | 9781 | 9781 | 4866.2 | 126.5 | 0.50 | 1.00 | 1.00 | 1.00 | -0.100 |
| Buonanno & Leonida, 2009 | 10 | -0.176 | 1600 | 1600 | 796.9 | 111.6 | 0.50 | 1.00 | 1.00 | 1.00 | -0.088 |
| Ou & Reynolds, 2010 | 10 | -0.239 | 374 | 359 | 118.9 | 62.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.179 |
| Machin et al., 2011 | 10 | -0.211 | 839 | 839 | 417.0 | 99.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.158 |
| Brugard & Falch, 2011 | 10 | -0.282 | 34914 | 16716 | 692.5 | 109.3 | 0.75 | 1.00 | 1.00 | 1.00 | -0.212 |
| Webbink et al., 2008 | 10 | -0.186 | 1125 | 1125 | 98.3 | 56.0 | 1.00 | 1.00 | 1.00 | 1.00 | -0.186 |
| Bjerk, 2011 | 10 | -0.293 | 1286 | 672 | 437.1 | 100.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.220 |
| Van Dorn, et al., 2011 | 10 | -0.158 | 28987 | 5666 | 528.0 | 104.2 | 0.75 | 1.00 | 1.00 | 1.00 | -0.119 |
| Fergusson & Horwood, 2000 | 11 | 0.283 | 262 | 749 | 192.6 | 10.8 | 1.00 | 1.00 | 0.50 | 1.00 | 0.141 |
| Lipsey et al., 1997 | 11 | 0.371 | 375 | 1500 | 103.6 | 10.3 | 0.50 | 1.00 | 1.00 | 1.00 | 0.185 |
| Elbogen & Johnson, 2009 | 11 | 0.150 | 7353 | 27300 | 435.2 | 11.1 | 0.75 | 1.00 | 0.50 | 1.00 | 0.056 |
| Carpenter, 2007 | 11 | 0.025 | 4600 | 4600 | 1562.3 | 11.3 | 1.00 | 1.00 | 0.50 | 1.00 | 0.013 |
| WSIPP analysis | 11 | 0.599 | 4431 | 33147 | 978.2 | 11.3 | 0.50 | 1.00 | 1.00 | 1.00 | 0.299 |
| Van Dorn, et al., 2011 | 11 | 0.145 | 3465 | 31188 | 86.2 | 10.1 | 0.75 | 1.00 | 1.00 | 1.00 | 0.109 |
| Zuvekas et al., 2005 | 12 | -0.068 | 2887 | 6933 | 484.5 | 4.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.034 |
| Mullahy & Sindelar, 1996 | 12 | -0.407 | 2381 | 21425 | 1294.3 | 4.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.204 |
| Terza, 2002 | 12 | -1.042 | 982 | 8840 | 487.5 | 4.8 | 0.50 | 1.00 | 1.00 | 0.50 | -0.260 |
| Chevrou-Severac & Jeanrenaud, 2002 | 12 | -0.613 | 216 | 7283 | 49.6 | 4.4 | 0.50 | 1.00 | 1.00 | 0.50 | -0.153 |
| Feng et al., 2001 | 12 | 0.044 | 647 | 7475 | 241.9 | 4.7 | 0.75 | 1.00 | 1.00 | 0.50 | 0.016 |
| Auld, 2002 | 12 | -0.602 | 982 | 8840 | 387.1 | 4.7 | 0.50 | 1.00 | 1.00 | 0.50 | -0.150 |
| MacDonald & Shields, 2004 | 12 | -0.217 | 664 | 5980 | 298.8 | 4.7 | 0.50 | 1.00 | 1.00 | 0.50 | -0.054 |
| Cook & Peters, 2005 | 12 | -0.036 | 624 | 7432 | 381.0 | 4.7 | 1.00 | 1.00 | 1.00 | 0.50 | -0.018 |
| Saffer & Dave, 2005 | 12 | -0.082 | 210 | 6790 | 125.9 | 4.6 | 1.00 | 1.00 | 1.00 | 0.50 | -0.041 |
| Johansson et al, 2007 | 12 | -1.161 | 453 | 4298 | 216.0 | 4.7 | 0.50 | 0.75 | 1.00 | 1.00 | -0.435 |
| French et al., 2011 | 12 | -0.211 | 3819 | 36917 | 963.0 | 4.8 | 0.75 | 1.00 | 1.00 | 1.00 | -0.158 |
| Sangchai, 2006 | 12 | -0.371 | 1689 | 16339 | 659.0 | 4.8 | 0.75 | 1.00 | 1.00 | 1.00 | -0.278 |
| Sangchai, 2006 | 12 | -1.265 | 2273 | 21552 | 1018.1 | 4.8 | 0.75 | 1.00 | 1.00 | 1.00 | -0.949 |
| Zarkin et al., 1998 | 13 | -0.004 | 442 | 11683 | 425.9 | 160.0 | 0.50 | 1.00 | 1.00 | 0.50 | -0.001 |
| Kenkel & Ribar, 1994 | 13 | -0.172 | 1742 | 5346 | 1310.4 | 214.4 | 1.00 | 1.00 | 1.00 | 1.00 | -0.172 |
| Bray, 2005 | 13 | -0.017 | 277 | 1572 | 235.7 | 122.8 | 1.00 | 1.00 | 1.00 | 0.50 | -0.009 |
| Jones & Richmond, 2006 | 13 | -0.051 | 798 | 2848 | 622.9 | 181.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.038 |
| Renna, 2008 | 13 | -0.080 | 578 | 3548 | 496.5 | 169.0 | 1.00 | 1.00 | 1.00 | 1.00 | -0.080 |
| Barrett, 2002 | 13 | -0.118 | 1104 | 4601 | 889.4 | 199.0 | 0.50 | 1.00 | 1.00 | 0.50 | -0.029 |
| Keng & Huffman, 2010 | 13 | -0.119 | 1393 | 2707 | 918.4 | 200.4 | 1.00 | 1.00 | 1.00 | 0.50 | -0.060 |
| Peters, 2004 | 13 | -0.048 | 1930 | 6842 | 1505.0 | 219.0 | 1.00 | 1.00 | 1.00 | 0.50 | -0.024 |
| Auld, 2005 | 13 | 0.104 | 362 | 3529 | 328.1 | 143.9 | 0.50 | 1.00 | 0.50 | 1.00 | 0.026 |
| Buchmueller & Zuvekas, 1998 | 14 | -0.220 | 449 | 1651 | 178.8 | 178.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.110 |
| Zuvekas et al., 2005 | 14 | -0.171 | 929 | 8089 | 226.9 | 226.9 | 0.50 | 1.00 | 1.00 | 1.00 | -0.086 |
| Alexandre & French, 2004 | 14 | -0.285 | 926 | 553 | 226.1 | 226.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.214 |
| French et al., 2001 | 14 | -0.271 | 379 | 9242 | 216.0 | 216.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.204 |
| WSIPP analysis | 14 | -0.176 | 990 | 37077 | 640.8 | 640.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.088 |
| Zuvekas et al., 2005 | 15 | 0.000 | 929 | 8089 | 833.2 | 4.4 | 0.50 | 1.00 | 1.00 | 1.00 | 0.000 |
| Ringel et al., 2006 | 15 | -0.687 | 71 | 721 | 63.4 | 4.1 | 0.50 | 1.00 | 1.00 | 1.00 | -0.343 |

Exhibit E3

| Individual Effect Sizes Used in the Meta-Analysis of Each Topic | | | | | | | | | | | |
|--|--------------|-------------------------|--------------------------------|-------------------------|---------------------------------------|--|---|-----------------------------|--|--|----------------------|
| Study (Date) | Topic Number | Study Results | | | | | Multiplicative Weights & Adjusted Effect Size | | | | |
| | | Un-adjusted effect size | Number in test condition group | Number in control group | Inverse variance weight-fixed effects | Inverse variance weight-random effects | Re-search design | General-izability of sample | Relevance of inde-pendent variable measure | Relevance of de-pendent variable measure | Adjusted effect size |
| Van Dorn, et al., 2011 | 16 | 0.250 | 3465 | 31188 | 99.8 | 99.8 | 0.75 | 1.00 | 1.00 | 1.00 | 0.187 |
| Bray et al., 2000 | 17 | -0.508 | 630 | 762 | 111.1 | 48.4 | 0.75 | 1.00 | 1.00 | 1.00 | -0.381 |
| Ellickson et al., 1998 | 17 | -0.074 | 860 | 3530 | 228.8 | 62.5 | 0.75 | 1.00 | 1.00 | 1.00 | -0.056 |
| Mensch & Kandel, 1988 | 17 | -0.177 | 7567 | 4094 | 1317.0 | 80.7 | 0.75 | 1.00 | 1.00 | 1.00 | -0.132 |
| Yamada et al., 1996 | 17 | -0.432 | 75 | 597 | 17.1 | 14.3 | 0.50 | 1.00 | 1.00 | 1.00 | -0.216 |
| Fergusson & Horwood, 1997 | 17 | -0.385 | 180 | 755 | 73.9 | 39.7 | 1.00 | 0.75 | 1.00 | 1.00 | -0.289 |
| Brook et al., 2002 | 17 | -0.217 | 100 | 1048 | 61.9 | 36.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.162 |
| McCaffrey et al., 2010 | 17 | -0.112 | 276 | 2482 | 27.8 | 21.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.084 |
| van Ours & Williams, 2009 | 17 | -0.198 | 5931 | 5862 | 1994.5 | 82.4 | 0.50 | 1.00 | 1.00 | 1.00 | -0.099 |
| Horwood et al., 2010 | 17 | -0.476 | 420 | 624 | 155.1 | 55.3 | 1.00 | 0.75 | 1.00 | 1.00 | -0.357 |
| Horwood et al., 2010 | 17 | -0.185 | 482 | 1036 | 121.0 | 50.2 | 1.00 | 1.00 | 1.00 | 1.00 | -0.185 |
| Horwood et al., 2010 | 17 | -0.480 | 1418 | 2176 | 337.6 | 68.5 | 1.00 | 1.00 | 1.00 | 1.00 | -0.480 |
| Ensminger et al., 1996 | 17 | -0.621 | 109 | 456 | 53.8 | 33.1 | 0.75 | 0.75 | 1.00 | 1.00 | -0.349 |
| WSIPP analysis | 17 | -0.152 | 10890 | 13748 | 2235.8 | 82.7 | 0.50 | 1.00 | 1.00 | 1.00 | -0.076 |
| Jofre-Bonet et al., 2005 | 18 | -0.024 | 31105 | 88778 | 12150.8 | 137.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.018 |
| Dastan, 2011 | 18 | -0.147 | 4203 | 7806 | 1411.0 | 126.7 | 0.75 | 1.00 | 1.00 | 1.00 | -0.110 |
| Anger & Kvasnicka, 2010 | 19 | 0.000 | 819 | 1149 | 478.2 | 41.9 | 0.50 | 1.00 | 1.00 | 1.00 | 0.000 |
| Auld, 2005 | 19 | -0.579 | 1280 | 2611 | 828.3 | 43.6 | 0.50 | 1.00 | 1.00 | 1.00 | -0.289 |
| Jofre-Bonet et al., 2005 | 19 | -0.026 | 31105 | 88778 | 23033.1 | 45.9 | 0.75 | 1.00 | 1.00 | 1.00 | -0.019 |
| Dastan, 2011 | 19 | -0.029 | 4203 | 7806 | 2731.8 | 45.2 | 0.75 | 1.00 | 1.00 | 1.00 | -0.022 |
| Braakmann, 2008 | 19 | -0.015 | 3611 | 8647 | 2547.2 | 45.2 | 0.50 | 1.00 | 1.00 | 1.00 | -0.008 |
| Goerge et al., 2008 | 20 | 0.238 | 96227 | 1771669 | 17341.2 | 17341.2 | 0.50 | 1.00 | 1.00 | 1.00 | 0.119 |
| Goerge et al., 2008 | 21 | 0.116 | 96227 | 1771669 | 5818.9 | 5818.9 | 0.50 | 1.00 | 1.00 | 1.00 | 0.058 |
| Fletcher & Wolfe, 2009b | 22 | 0.137 | 564 | 149 | 35.0 | 13.6 | 0.75 | 1.00 | 1.00 | 1.00 | 0.103 |
| Hoffman, 2008 | 22 | 0.091 | 762 | 69 | 63.3 | 16.5 | 0.75 | 1.00 | 1.00 | 1.00 | 0.068 |
| Boden et al., 2008 | 22 | 0.820 | 22 | 429 | 10.0 | 6.9 | 0.50 | 1.00 | 1.00 | 1.00 | 0.410 |
| Angrist & Lavy, 1996 | 23 | 0.213 | 557 | 17238 | 539.2 | 539.2 | 1.00 | 1.00 | 1.00 | 1.00 | 0.213 |
| Angrist & Lavy, 1996 | 23 | 0.148 | 500 | 541 | 259.1 | 259.1 | 1.00 | 1.00 | 1.00 | 1.00 | 0.148 |
| Moore et al., 1997 | 23 | 0.245 | 77 | 199 | 24.2 | 24.2 | 0.50 | 1.00 | 1.00 | 1.00 | 0.123 |
| Levine et al., 2007 | 23 | 0.320 | 451 | 354 | 84.0 | 84.0 | 1.00 | 1.00 | 1.00 | 1.00 | 0.320 |
| Hoffman & Scher, 2008 | 24 | -0.205 | 644 | 337 | 86.8 | 86.8 | 0.75 | 1.00 | 1.00 | 1.00 | -0.154 |
| Manlove et al., 2008 | 24 | -0.150 | 221 | 461 | 73.8 | 73.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.075 |
| Francesconi, 2008 | 24 | -0.314 | 85 | 1098 | 53.6 | 53.6 | 1.00 | 1.00 | 1.00 | 0.50 | -0.157 |
| Fletcher & Wolfe, 2009b | 25 | -0.241 | 563 | 148 | 71.1 | 71.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.181 |
| Fletcher, 2010 | 25 | -0.192 | 233 | 2094 | 68.9 | 68.9 | 1.00 | 1.00 | 1.00 | 1.00 | -0.192 |
| Webbink et al., 2009 | 25 | -0.065 | 77 | 77 | 25.4 | 25.4 | 1.00 | 1.00 | 1.00 | 1.00 | -0.065 |
| Hoffman, 2008 | 25 | -0.096 | 453 | 41 | 25.3 | 25.3 | 0.75 | 1.00 | 1.00 | 1.00 | -0.072 |
| Pogarsky et al., 2003 | 26 | 0.145 | 228 | 457 | 151.9 | 151.9 | 0.75 | 1.00 | 1.00 | 1.00 | 0.109 |
| Scher & Hoffman, 2008 | 26 | 0.268 | 465 | 1158 | 67.6 | 67.6 | 0.75 | 1.00 | 1.00 | 1.00 | 0.201 |
| Ettner et al., 1997 | 27 | -0.570 | 1327 | 3299 | 388.0 | 16.8 | 0.75 | 1.00 | 1.00 | 1.00 | -0.428 |
| Farahati et al., 2003 | 27 | -0.255 | 74 | 438 | 32.8 | 11.5 | 0.75 | 1.00 | 1.00 | 1.00 | -0.191 |
| Savoca & Rosenheck, 2000 | 27 | -0.624 | 315 | 1102 | 115.9 | 15.3 | 0.50 | 0.75 | 1.00 | 1.00 | -0.234 |
| Alexandre & French, 2001 | 27 | -0.527 | 384 | 890 | 144.5 | 15.7 | 0.50 | 0.75 | 1.00 | 1.00 | -0.198 |
| Chatterji et al., 2007 | 27 | -0.331 | 535 | 3538 | 302.8 | 16.6 | 0.50 | 1.00 | 1.00 | 1.00 | -0.165 |
| Cornwell et al., 2009 | 27 | -0.059 | 1852 | 8790 | 735.0 | 17.2 | 0.75 | 1.00 | 1.00 | 1.00 | -0.044 |
| Ojeda et al., 2010 | 27 | -0.881 | 2805 | 27418 | 1351.7 | 17.4 | 0.75 | 1.00 | 0.50 | 1.00 | -0.330 |

Exhibit E3

| Individual Effect Sizes Used in the Meta-Analysis of Each Topic | | | | | | | | | | | |
|--|--------------|-------------------------|--------------------------------|-------------------------|---------------------------------------|--|---|-----------------------------|--|--|----------------------|
| Study (Date) | Topic Number | Study Results | | | | | Multiplicative Weights & Adjusted Effect Size | | | | |
| | | Un-adjusted effect size | Number in test condition group | Number in control group | Inverse variance weight-fixed effects | Inverse variance weight-random effects | Re-search design | General-izability of sample | Relevance of inde-pendent variable measure | Relevance of de-pendent variable measure | Adjusted effect size |
| Gibb et al., 2010 | 27 | -0.251 | 238 | 713 | 60.9 | 13.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.188 |
| Cowell et al., 2009 | 27 | -0.260 | 4749 | 28326 | 2718.1 | 17.5 | 0.50 | 1.00 | 1.00 | 1.00 | -0.130 |
| Alexandre et al., 2004 | 27 | -0.204 | 1038 | 14371 | 145.7 | 15.7 | 0.50 | 1.00 | 1.00 | 1.00 | -0.102 |
| Frijters et al., 2010 | 27 | -0.901 | 1716 | 5946 | 804.3 | 17.2 | 1.00 | 1.00 | 0.50 | 1.00 | -0.451 |
| Bruffaerts et al., 2009 | 27 | -0.690 | 42 | 821 | 26.5 | 10.6 | 0.50 | 1.00 | 1.00 | 1.00 | -0.345 |
| Zhang et al., 2009 | 27 | -0.251 | 4252 | 26040 | 1736.1 | 17.4 | 0.75 | 1.00 | 0.50 | 1.00 | -0.094 |
| Chatterji et al., 2009 | 27 | -0.169 | 2536 | 9277 | 1081.7 | 17.3 | 0.50 | 1.00 | 1.00 | 1.00 | -0.085 |
| Tian et al., 2005 | 27 | -0.150 | 459 | 5239 | 280.0 | 16.5 | 0.50 | 0.75 | 1.00 | 1.00 | -0.056 |
| Baldwin & Marcus, 2007 | 27 | -0.577 | 1149 | 9675 | 456.4 | 16.9 | 0.50 | 1.00 | 1.00 | 1.00 | -0.288 |
| Jofre-Bonet et al., 2005 | 27 | -0.457 | 13251 | 106632 | 7517.7 | 17.5 | 0.75 | 1.00 | 0.50 | 1.00 | -0.171 |
| WSIPP analysis | 27 | -0.197 | 1865 | 36202 | 1177.5 | 17.3 | 0.50 | 1.00 | 1.00 | 1.00 | -0.098 |
| Ettner et al., 1997 | 28 | -0.212 | 1327 | 3299 | 942.2 | 316.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.159 |
| Marcotte & Wilcox-Gök, 2003 | 28 | -0.164 | 1029 | 2402 | 718.5 | 286.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.123 |
| Frank & Gertler, 1991 | 28 | -0.281 | 106 | 776 | 92.8 | 77.6 | 0.50 | 1.00 | 1.00 | 1.00 | -0.141 |
| French & Zarkin, 1998 | 28 | -0.124 | 45 | 363 | 39.9 | 36.8 | 0.50 | 0.75 | 0.50 | 1.00 | -0.023 |
| Baldwin & Marcus, 2007 | 28 | -0.051 | 1149 | 9675 | 1026.9 | 324.9 | 0.50 | 1.00 | 1.00 | 1.00 | -0.025 |
| Cseh, 2008 | 28 | -0.039 | 1379 | 5657 | 1108.7 | 332.7 | 1.00 | 1.00 | 1.00 | 1.00 | -0.039 |
| Forbes et al., 2010 | 28 | -0.042 | 5843 | 20959 | 4568.4 | 430.6 | 0.50 | 1.00 | 1.00 | 1.00 | -0.021 |
| Jofre-Bonet et al., 2005 | 28 | -0.088 | 13251 | 106632 | 11782.2 | 456.9 | 0.75 | 1.00 | 0.50 | 1.00 | -0.033 |
| Chatterji et al., 2011 | 28 | -0.052 | 911 | 3224 | 710.3 | 284.8 | 0.75 | 1.00 | 1.00 | 1.00 | -0.039 |
| Ettner et al., 1997 | 29 | -0.328 | 454 | 4172 | 170.6 | 29.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.164 |
| Farahati et al., 2003 | 29 | -0.255 | 74 | 438 | 32.8 | 17.2 | 0.75 | 1.00 | 1.00 | 1.00 | -0.191 |
| Savoca & Rosenheck, 2000 | 29 | -0.602 | 79 | 1338 | 39.7 | 18.9 | 0.50 | 0.75 | 1.00 | 1.00 | -0.226 |
| Alexandre & French, 2001 | 29 | -0.527 | 384 | 890 | 144.5 | 28.9 | 0.50 | 0.75 | 1.00 | 1.00 | -0.198 |
| Cornwell et al., 2009 | 29 | -0.099 | 724 | 9917 | 336.8 | 32.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.075 |
| Gibb et al., 2010 | 29 | -0.351 | 143 | 808 | 46.6 | 20.4 | 0.75 | 1.00 | 1.00 | 1.00 | -0.263 |
| Cowell et al., 2009 | 29 | -0.270 | 1534 | 31541 | 987.0 | 34.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.135 |
| Chatterji et al., 2009 | 29 | -0.310 | 1709 | 10104 | 861.7 | 34.7 | 0.50 | 1.00 | 1.00 | 1.00 | -0.155 |
| Tian et al., 2005 | 29 | -0.150 | 459 | 5239 | 280.0 | 32.0 | 0.50 | 0.75 | 1.00 | 1.00 | -0.056 |
| Baldwin & Marcus, 2007 | 29 | -0.687 | 703 | 10121 | 327.5 | 32.5 | 0.50 | 1.00 | 1.00 | 1.00 | -0.343 |
| Ettner et al., 1997 | 30 | -0.086 | 454 | 4172 | 409.0 | 409.0 | 0.50 | 1.00 | 1.00 | 1.00 | -0.043 |
| Marcotte & Wilcox-Gök, 2003 | 30 | 0.008 | 483 | 2948 | 415.2 | 415.2 | 0.75 | 1.00 | 1.00 | 1.00 | 0.006 |
| Baldwin & Marcus, 2007 | 30 | -0.055 | 703 | 10121 | 657.3 | 657.3 | 0.50 | 1.00 | 1.00 | 1.00 | -0.027 |
| Cseh, 2008 | 30 | -0.039 | 1379 | 5657 | 1108.7 | 1108.7 | 1.00 | 1.00 | 1.00 | 1.00 | -0.039 |
| Ettner et al., 1997 | 31 | -0.087 | 562 | 4064 | 163.1 | 31.2 | 0.50 | 1.00 | 1.00 | 1.00 | -0.043 |
| Savoca & Rosenheck, 2000 | 31 | -0.632 | 235 | 1182 | 97.1 | 27.6 | 0.50 | 0.75 | 1.00 | 1.00 | -0.237 |
| Cornwell et al., 2009 | 31 | -0.032 | 1128 | 9513 | 476.1 | 35.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.024 |
| Gibb et al., 2010 | 31 | -0.151 | 143 | 808 | 39.0 | 19.4 | 0.75 | 1.00 | 1.00 | 1.00 | -0.114 |
| Cowell et al., 2009 | 31 | -0.116 | 2301 | 30774 | 1395.3 | 37.5 | 0.50 | 1.00 | 1.00 | 1.00 | -0.058 |
| Chatterji et al., 2009 | 31 | -0.117 | 1168 | 10645 | 561.8 | 36.0 | 0.50 | 1.00 | 1.00 | 1.00 | -0.058 |
| Baldwin & Marcus, 2007 | 31 | -0.583 | 294 | 10530 | 136.7 | 30.0 | 0.50 | 1.00 | 1.00 | 1.00 | -0.292 |
| Ettner et al., 1997 | 32 | -0.029 | 562 | 4064 | 493.8 | 25.0 | 0.50 | 1.00 | 1.00 | 1.00 | -0.014 |
| Marcotte & Wilcox-Gök, 2003 | 32 | -0.385 | 752 | 2679 | 579.7 | 25.2 | 0.75 | 1.00 | 1.00 | 1.00 | -0.289 |
| Baldwin & Marcus, 2007 | 32 | -0.143 | 294 | 10530 | 285.9 | 24.1 | 0.50 | 1.00 | 1.00 | 1.00 | -0.072 |
| Fergusson & Lynskey, 1998 | 33 | 0.465 | 83 | 886 | 13.8 | 11.7 | 0.50 | 1.00 | 1.00 | 1.00 | 0.232 |
| Satterfield et al., 2007 | 33 | 0.678 | 169 | 64 | 25.0 | 18.9 | 0.50 | 1.00 | 1.00 | 1.00 | 0.339 |

Exhibit E3

| Individual Effect Sizes Used in the Meta-Analysis of Each Topic | | | | | | | | | | | |
|--|--------------|-------------------------|--------------------------------|-------------------------|---------------------------------------|--|---|-----------------------------|--|--|----------------------|
| Study (Date) | Topic Number | Study Results | | | | | Multiplicative Weights & Adjusted Effect Size | | | | |
| | | Un-adjusted effect size | Number in test condition group | Number in control group | Inverse variance weight-fixed effects | Inverse variance weight-random effects | Re-search design | General-izability of sample | Relevance of inde-pendent variable measure | Relevance of de-pendent variable measure | Adjusted effect size |
| Fletcher & Wolfe, 2009a | 33 | 0.388 | 691 | 2947 | 303.0 | 61.1 | 0.50 | 1.00 | 1.00 | 1.00 | 0.194 |
| Copeland et al., 2007 | 33 | 0.339 | 125 | 1296 | 44.9 | 28.3 | 0.75 | 1.00 | 1.00 | 1.00 | 0.254 |
| Fergusson et al., 2005 | 33 | 0.763 | 46 | 927 | 17.4 | 14.2 | 0.75 | 1.00 | 1.00 | 1.00 | 0.572 |
| Bussing et al., 2010 | 33 | 0.684 | 94 | 163 | 10.3 | 9.1 | 0.75 | 1.00 | 1.00 | 0.50 | 0.256 |
| Murray et al., 2010 | 33 | 0.360 | 1090 | 7296 | 427.4 | 64.9 | 0.75 | 1.00 | 1.00 | 1.00 | 0.270 |
| Currie & Stabile, 2009 | 33 | 0.192 | 164 | 3056 | 105.5 | 44.4 | 0.75 | 1.00 | 1.00 | 1.00 | 0.144 |
| Currie & Stabile, 2009 | 33 | 0.364 | 116 | 2162 | 74.8 | 37.8 | 0.75 | 1.00 | 1.00 | 1.00 | 0.273 |
| Currie & Stabile, 2009 | 33 | 0.101 | 323 | 2903 | 197.6 | 55.2 | 0.75 | 1.00 | 1.00 | 1.00 | 0.075 |
| Currie & Stabile, 2009 | 33 | 0.163 | 228 | 2050 | 135.5 | 48.9 | 0.75 | 1.00 | 1.00 | 1.00 | 0.122 |
| Webbink et al., 2011 | 33 | 0.456 | 239 | 1899 | 62.3 | 34.3 | 1.00 | 1.00 | 1.00 | 1.00 | 0.456 |
| Fletcher & Wolfe, 2008 | 34 | -0.305 | 262 | 2645 | 93.8 | 62.2 | 0.50 | 1.00 | 1.00 | 1.00 | -0.152 |
| McLeod & Kaiser, 2004 | 34 | -0.273 | 57 | 367 | 22.0 | 19.7 | 0.75 | 1.00 | 1.00 | 1.00 | -0.205 |
| Fergusson & Lynskey, 1998 | 34 | -0.333 | 83 | 886 | 41.0 | 33.5 | 0.50 | 1.00 | 1.00 | 1.00 | -0.167 |
| Breslau et al., 2008 | 34 | -0.555 | 380 | 5206 | 191.3 | 93.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.278 |
| Breslau et al., 2008 | 34 | -0.322 | 486 | 5100 | 201.0 | 96.1 | 0.50 | 1.00 | 1.00 | 1.00 | -0.161 |
| Breslau et al., 2008 | 34 | -0.389 | 704 | 4882 | 288.5 | 112.4 | 0.50 | 1.00 | 1.00 | 1.00 | -0.194 |
| Currie et al., 2010 | 34 | -0.234 | 1739 | 48665 | 1090.5 | 157.5 | 1.00 | 1.00 | 1.00 | 1.00 | -0.234 |
| Galera et al., 2009 | 34 | -0.369 | 71 | 643 | 22.6 | 20.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.277 |
| Galera et al., 2009 | 34 | -0.438 | 71 | 643 | 21.0 | 18.8 | 0.75 | 1.00 | 1.00 | 1.00 | -0.328 |
| Currie & Stabile, 2009 | 34 | -0.217 | 127 | 2355 | 55.5 | 42.7 | 0.75 | 1.00 | 1.00 | 1.00 | -0.162 |
| Currie & Stabile, 2009 | 34 | -0.240 | 249 | 2237 | 104.1 | 66.5 | 0.75 | 1.00 | 1.00 | 1.00 | -0.180 |
| Currie & Stabile, 2009 | 34 | -0.552 | 132 | 2466 | 67.1 | 49.2 | 0.75 | 1.00 | 1.00 | 1.00 | -0.414 |
| Currie & Stabile, 2009 | 34 | -0.184 | 260 | 2339 | 90.7 | 60.8 | 0.75 | 1.00 | 1.00 | 1.00 | -0.138 |
| Breslau et al., 2011 | 34 | -0.386 | 1513 | 28149 | 767.8 | 148.5 | 0.75 | 1.00 | 1.00 | 1.00 | -0.289 |
| Breslau et al., 2011 | 34 | -0.309 | 2966 | 26696 | 1328.6 | 161.7 | 0.75 | 1.00 | 1.00 | 1.00 | -0.232 |
| Webbink et al., 2011 | 34 | -0.233 | 216 | 1712 | 103.0 | 66.1 | 1.00 | 1.00 | 1.00 | 1.00 | -0.233 |
| Porche et al., 2011 | 34 | -0.525 | 287 | 2245 | 129.6 | 76.1 | 0.50 | 1.00 | 1.00 | 1.00 | -0.263 |
| McLeod & Kaiser, 2004 | 35 | -0.210 | 75 | 349 | 26.0 | 22.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.158 |
| Duchesne et al., 2008 | 35 | -0.215 | 177 | 1640 | 93.6 | 60.9 | 0.75 | 1.00 | 0.50 | 1.00 | -0.081 |
| Breslau et al., 2008 | 35 | -0.303 | 654 | 4932 | 254.8 | 103.5 | 0.50 | 1.00 | 1.00 | 1.00 | -0.152 |
| Breslau et al., 2008 | 35 | -0.159 | 1782 | 3804 | 455.9 | 126.1 | 0.50 | 1.00 | 1.00 | 1.00 | -0.079 |
| Fletcher, 2010 | 35 | -0.167 | 186 | 2141 | 55.2 | 41.9 | 1.00 | 1.00 | 1.00 | 1.00 | -0.167 |
| Fergusson & Woodward, 2002 | 35 | -0.058 | 124 | 840 | 43.4 | 34.8 | 0.75 | 1.00 | 1.00 | 0.50 | -0.022 |
| Needham, 2009 | 35 | -0.140 | 1564 | 12657 | 365.6 | 118.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.105 |
| Currie & Stabile, 2009 | 35 | 0.055 | 443 | 2043 | 129.7 | 74.4 | 0.75 | 1.00 | 1.00 | 1.00 | 0.041 |
| Currie & Stabile, 2009 | 35 | -0.029 | 463 | 2140 | 126.2 | 73.2 | 0.75 | 1.00 | 1.00 | 1.00 | -0.022 |
| Breslau et al., 2011 | 35 | -0.054 | 7207 | 22455 | 2180.6 | 161.4 | 0.75 | 1.00 | 1.00 | 1.00 | -0.041 |
| Porche et al., 2011 | 35 | 0.018 | 368 | 2164 | 111.4 | 68.0 | 0.50 | 1.00 | 1.00 | 1.00 | 0.009 |
| Fletcher & Wolfe, 2008 | 36 | 0.453 | 261 | 2643 | 54.0 | 53.8 | 0.50 | 1.00 | 0.50 | 1.00 | 0.113 |
| Galera et al., 2009 | 36 | 0.597 | 163 | 1101 | 92.8 | 92.2 | 0.75 | 1.00 | 1.00 | 1.00 | 0.448 |
| Currie & Stabile, 2009 | 36 | 0.347 | 359 | 3232 | 99.3 | 98.6 | 0.75 | 1.00 | 1.00 | 1.00 | 0.260 |
| Currie & Stabile, 2009 | 36 | 0.468 | 582 | 5240 | 179.4 | 177.2 | 0.75 | 1.00 | 1.00 | 1.00 | 0.351 |
| Currie & Stabile, 2009 | 37 | -0.283 | 258 | 2318 | 230.9 | 31.9 | 0.75 | 1.00 | 1.00 | 1.00 | -0.212 |
| Currie & Stabile, 2009 | 37 | -0.262 | 258 | 2318 | 231.0 | 32.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.197 |
| Currie & Stabile, 2009 | 37 | -0.584 | 238 | 2142 | 211.0 | 31.5 | 0.75 | 1.00 | 1.00 | 1.00 | -0.438 |
| Fletcher & Wolfe, 2008 | 38 | -0.305 | 262 | 2645 | 93.8 | 93.8 | 0.50 | 1.00 | 1.00 | 1.00 | -0.152 |

Exhibit E3

| Individual Effect Sizes Used in the Meta-Analysis of Each Topic | | | | | | | | | | | |
|--|--------------|-------------------------|--------------------------------|-------------------------|---------------------------------------|--|---|-----------------------------|---|---|----------------------|
| Study (Date) | Topic Number | Study Results | | | | | Multiplicative Weights & Adjusted Effect Size | | | | |
| | | Un-adjusted effect size | Number in test condition group | Number in control group | Inverse variance weight-fixed effects | Inverse variance weight-random effects | Re-search design | General-izability of sample | Relevance of independent variable measure | Relevance of dependent variable measure | Adjusted effect size |
| Breslau et al., 2008 | 38 | -0.322 | 486 | 5100 | 201.0 | 201.0 | 0.50 | 1.00 | 1.00 | 1.00 | -0.161 |
| Galera et al., 2009 | 38 | -0.369 | 71 | 643 | 22.6 | 22.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.277 |
| Currie & Stabile, 2009 | 38 | -0.240 | 249 | 2237 | 104.1 | 104.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.180 |
| Currie & Stabile, 2009 | 38 | -0.184 | 260 | 2339 | 90.7 | 90.7 | 0.75 | 1.00 | 1.00 | 1.00 | -0.138 |
| Breslau et al., 2011 | 38 | -0.309 | 2966 | 26696 | 1328.6 | 1328.6 | 0.75 | 1.00 | 1.00 | 1.00 | -0.232 |
| Satterfield et al., 2007 | 39 | 0.678 | 169 | 64 | 25.0 | 14.3 | 0.50 | 1.00 | 1.00 | 1.00 | 0.339 |
| Fletcher & Wolfe, 2009a | 39 | 0.388 | 691 | 2947 | 303.0 | 30.2 | 0.50 | 1.00 | 1.00 | 1.00 | 0.194 |
| Bussing et al., 2010 | 39 | 0.684 | 94 | 163 | 10.3 | 7.9 | 0.75 | 1.00 | 1.00 | 0.50 | 0.256 |
| Currie & Stabile, 2009 | 39 | 0.101 | 323 | 2903 | 197.6 | 28.7 | 0.75 | 1.00 | 1.00 | 1.00 | 0.075 |
| Currie & Stabile, 2009 | 39 | 0.163 | 228 | 2050 | 135.5 | 26.9 | 0.75 | 1.00 | 1.00 | 1.00 | 0.122 |
| Fergusson & Lynskey, 1998 | 40 | -0.333 | 83 | 886 | 41.0 | 34.4 | 0.50 | 1.00 | 1.00 | 1.00 | -0.167 |
| Breslau et al., 2008 | 40 | -0.555 | 380 | 5206 | 191.3 | 100.7 | 0.50 | 1.00 | 1.00 | 1.00 | -0.278 |
| Galera et al., 2009 | 40 | -0.438 | 71 | 643 | 21.0 | 19.1 | 0.75 | 1.00 | 1.00 | 1.00 | -0.328 |
| Currie & Stabile, 2009 | 40 | -0.217 | 127 | 2355 | 55.5 | 44.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.162 |
| Currie & Stabile, 2009 | 40 | -0.552 | 132 | 2466 | 67.1 | 51.0 | 0.75 | 1.00 | 1.00 | 1.00 | -0.414 |
| Breslau et al., 2011 | 40 | -0.386 | 1513 | 28149 | 767.8 | 166.5 | 0.75 | 1.00 | 1.00 | 1.00 | -0.289 |
| Porche et al., 2011 | 40 | -0.525 | 287 | 2245 | 129.6 | 80.5 | 0.50 | 1.00 | 1.00 | 1.00 | -0.263 |
| Fergusson & Lynskey, 1998 | 41 | 0.465 | 83 | 886 | 13.8 | 13.6 | 0.50 | 1.00 | 1.00 | 1.00 | 0.232 |
| Copeland et al., 2007 | 41 | 0.339 | 125 | 1296 | 44.9 | 42.7 | 0.75 | 1.00 | 1.00 | 1.00 | 0.254 |
| Fergusson et al., 2005 | 41 | 0.763 | 46 | 927 | 17.4 | 17.1 | 0.75 | 1.00 | 1.00 | 1.00 | 0.572 |
| Murray et al., 2010 | 41 | 0.360 | 1090 | 7296 | 427.4 | 286.6 | 0.75 | 1.00 | 1.00 | 1.00 | 0.270 |
| Currie & Stabile, 2009 | 41 | 0.192 | 164 | 3056 | 105.5 | 94.1 | 0.75 | 1.00 | 1.00 | 1.00 | 0.144 |
| Currie & Stabile, 2009 | 41 | 0.364 | 116 | 2162 | 74.8 | 68.8 | 0.75 | 1.00 | 1.00 | 1.00 | 0.273 |
| Webbink et al., 2011 | 41 | 0.456 | 239 | 1899 | 62.3 | 58.1 | 1.00 | 1.00 | 1.00 | 1.00 | 0.456 |
| Steadman et al., 1998 | 42 | 0.340 | 286 | 519 | 29.5 | 4.4 | 0.50 | 1.00 | 1.00 | 1.00 | 0.170 |
| Fazel et al., 2010 | 42 | 0.938 | 2570 | 4059 | 426.6 | 5.1 | 0.50 | 1.00 | 1.00 | 1.00 | 0.469 |
| Fazel et al., 2009 | 42 | 0.285 | 4680 | 8118 | 710.8 | 5.1 | 0.50 | 1.00 | 1.00 | 1.00 | 0.142 |

Studies used in the links meta-analyses

- Alexandre, P. K., & French, M. T. (2001). Labor supply of poor residents in metropolitan Miami, Florida: The role of depression and the co-morbid effects of substance use. *The Journal of Mental Health Policy and Economics*, 4(4), 161-173.
- Alexandre, P. K., & French, M. T. (2004). Further evidence on the labor market effects of addiction: Chronic drug use and labor supply in metropolitan Miami. *Contemporary Economic Policy*, 22(3), 382-393.
- Alexandre, P. K., Fede, J. Y., & Mullings, M. (2004). *Gender differences in the labor market effects of serious mental illness*. In D. E. Marcotte (Ed.), *The economics of gender and mental illness* (Research in Human Capital and Development, Vol. 15; pp. 53-71). United Kingdom: Emerald Group Publishing.
- Anger, S., & Kvasnicka, M. (2010). Does smoking really harm your earnings so much? Biases in current estimates of the smoking wage penalty. *Applied Economics Letters*, 17(6), 561-564.
- Angrist, J. D., & Lavy, V. (1996, October). *The effect of teen childbearing and single parenthood on childhood disabilities and progress in school* (Working Paper No. 5807). Cambridge, MA: National Bureau of Economic Research.
- Apel, R., & Sweeten, G. (2009, September). *The effect of criminal justice involvement in the transition to adulthood* (Document No. NCJ 228380). Washington, DC: National Institute of Justice.
- Arteaga, I., Chen, C. C., & Reynolds, A. J. (2010). Childhood predictors of adult substance abuse. *Children and Youth Services Review*, 32(8), 1108-1120.
- Auld, M. C. (2002, November). *Robust system estimation of causal effects on binary outcomes, with application to effect of alcohol abuse on employment*. Calgary, AB, Canada: University of Calgary. Retrieved June 6, 2011, from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.7.6712>
- Auld, M. C. (2005). Smoking, drinking, and income. *Journal of Human Resources*, 40(2), 505-519.
- Baldwin, M. L., & Marcus, S. C. (2007). Labor market outcomes of persons with mental disorders. *Industrial Relations*, 46(3), 481-510.
- Barrett, G. F. (2002). The effect of alcohol consumption on earnings. *The Economic Record*, 78(240), 79-96.
- Bjerk, D. (2011, September). *Re-examining the impact of dropping out on criminal and labor outcomes in early adulthood* (IZA Discussion Paper No. 5995). Bonn, Germany: Institute for the Study of Labor. Retrieved from <http://ftp.iza.org/dp5995.pdf>
- Boden, J. M., Fergusson, D. M., & John, H. L. (2008). Early motherhood and subsequent life outcomes. *Journal of Child Psychology and Psychiatry*, 49(2), 151-160.
- Boden, J. M., Horwood, L. J., & Fergusson, D. M. (2007). Exposure to childhood sexual and physical abuse and subsequent educational achievement outcomes. *Child Abuse & Neglect*, 31(10), 1101-1114.
- Braakmann, N. (2008, August). *The smoking wage penalty in the United Kingdom: Regression and matching evidence from the British Household Panel Survey* (Working Paper Series in Economics No. 96). Luneburg, Germany: University of Luneburg.
- Bray, J. W. (2005). Alcohol use, human capital, and wages. *Journal of Labor Economics*, 23(2), 279-312.
- Bray, J. W., Zarkin, G. A., Ringwalt, C., & Qi, J. (2000). The relationship between marijuana initiation and dropping out of high school. *Health Economics*, 9(1), 9-18.
- Breslau, J., Lane, M., Sampson, N., & Kessler, R. C. (2008). Mental disorders and subsequent educational attainment in a US national sample. *Journal of Psychiatric Research*, 42(9), 708-716.
- Breslau, J., Miller, E., Joanie, C. W. J., & Schweitzer, J. B. (2011). Childhood and adolescent onset psychiatric disorders, substance use, and failure to graduate high school on time. *Journal of Psychiatric Research*, 45(3), 295-301.
- Brook, J. S., Adams, R. E., Balka, E. B., & Johnson, E. (2002). Early adolescent marijuana use: Risks for the transition to young adulthood. *Psychological Medicine*, 32(1), 79-91.
- Bruffaerts, R., Bonnewyn, A., & Demyttenaere, K. (2009). The individual and societal effects of non-psychotic serious mental disorders on earnings in Belgium. *European Psychiatry*, 24(4), 207-213.
- Brugard, K. H., & Falch, T. (2011, March). *High school attainment, student achievement, and imprisonment*. Unpublished manuscript, Norwegian University of Science and Technology.
- Buchmueller, T. C., & Zuvekas, S. H. (1998). Drug use, drug abuse, and labour market outcomes. *Health Economics*, 7(3), 229-245.
- Buonanno, P., & Leonida, L. (2009). Non-market effects of education on crime: Evidence from Italian regions. *Economics of Education Review*, 28(1), 11-17.
- Bussing, R., Mason, D. M., Bell, L., Porter, P., & Garvan, C. (2010). Adolescent outcomes of childhood Attention-Deficit/Hyperactivity Disorder in a diverse community sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(6), 595-605.
- Carpenter, C. (2007). Heavy alcohol use and crime: Evidence from underage drunk-driving laws. *The Journal of Law & Economics*, 50(3), 539-557.
- Chapman, D. P., Whitfield, C. L., Felitti, V. J., Dube, S. R., Edwards, V. J., & Anda, R. F. (2004). Adverse childhood experiences and the risk of depressive disorders in adulthood. *Journal of Affective Disorders*, 82(2), 217-225.
- Chatterji, P., Alegria, M., Lu, M., & Takeuchi, D. (2007). Psychiatric disorders and labor market outcomes: Evidence from the National Latino and Asian American Study. *Health Economics*, 16(10), 1069-1090.
- Chatterji, P., Alegria, M., & Takeuchi, D. (2009). Racial/ethnic differences in the effects of psychiatric disorders on employment. *Atlantic Economic Journal*, 37(3), 243-257.
- Chatterji, P., Alegria, M., & Takeuchi, D. (2011). Psychiatric disorders and labor market outcomes: Evidence from the National Comorbidity Survey-Replication. *Journal of Health Economics*, 30(5), 858-868.
- Chevrou-Severac, H., & Jeanrenaud, C. (2002). *The impact of alcohol abuse on employment in Switzerland*. Neuchâtel, Switzerland: Institut de Recherche Economique et Régional. Retrieved June 6, 2011, from <http://perso.wanadoo.fr/ces/Pages/english/Poster17.pdf>
- Cohen, P., Smailes, E., & Brown, J. (2004). Effects of childhood maltreatment on adult arrests in a general population sample (Document No. 199707). In B. S. Fisher (Ed.), *Violence against women and family violence: Developments in research, practice, and policy* (pp. II-1-1-II-1-10). Washington, DC: National Institute of Justice.
- Cook, P. J., & Peters, B. (2005, December). *The myth of the drinker's bonus* (Working Paper No. 11902). Cambridge: National Bureau of Economic Research.
- Copeland, W. E., Miller-Johnson, S., Keeler, G., Angold, A., & Costello, E. J. (2007). Childhood psychiatric disorders and young adult crime: A prospective, population-based study. *The American Journal of Psychiatry*, 164(11), 1668-1675.
- Cornwell, K., Forbes, C., Inder, B., & Meadows, G. (2009). Mental illness and its effects on labour market outcomes. *The Journal of Mental Health Policy and Economics*, 12(3), 107-118.

Studies used in the links meta-analyses

- Cowell, A. J., Luo, Z., & Masuda, Y. J. (2009). Psychiatric disorders and the labor market: An analysis by disorder profiles. *The Journal of Mental Health Policy and Economics*, 12(1), 3-17.
- Cseh, A. (2008). The effects of depressive symptoms on earnings. *Southern Economic Journal*, 75(2), 383-409.
- Currie, J., & Stabile, M. (2006). Child mental health and human capital accumulation: The case of ADHD. *Journal of Health Economics*, 25(6), 1094-1118.
- Currie, J., Stabile, M., Manivong, P., & Roos, L. L. (2010). Child health and young adult outcomes. *Journal of Human Resources*, 45(3), 517-548.
- Currie, J., & Tekin, E. (2006, June). *Does child abuse cause crime?* (Working Paper No. W12171). Cambridge: National Bureau of Economic Research.
- Dastan, I. (2011). Labor market effects of obesity, smoking, and alcohol use. *Dissertation Abstracts International*, 72(03), A.
- Duchesne, S., Vitaro, F., Larose, S., & Tremblay, R. E. (2008). Trajectories of anxiety during elementary-school years and the prediction of high school noncompletion. *Journal of Youth and Adolescence*, 37(9), 1134-1146.
- Elbogen, E. B., & Johnson, S. C. (2009). The intricate link between violence and mental disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, 66(2), 152-161.
- Ellickson, P. E., Bui, K., Bell, R., & McGuigan, K. A. (1998). Does early drug use increase the risk of dropping out of high school? *Journal of Drug Issues*, 28(2), 357-380.
- English, D. J., Widom, C. S., & Brandford, C. (2002). *Childhood victimization and delinquency, adult criminality, and violent criminal behavior: A replication and extension* (Document No. 192291). Seattle, WA: Office of Children's Administration Research.
- Ensminger, M. E., Lamkin, R. P., & Jacobson, N. (1996). School leaving: A longitudinal perspective including neighborhood effects. *Child Development*, 67(5), 2400-2416.
- Ettner, S. L., Frank, R. G., & Kessler, R. C. (1997). The impact of psychiatric disorders on labor market outcomes. *Industrial & Labour Relations Review*, 51(1), 64-81.
- Farahati, F., Booth, B., & Wilcox-Gök, V. (2003, April). *Employment effects of comorbid depression and substance use* (Working Paper). North Little Rock, AR: University of Arkansas for Medical Sciences, Centers for Mental Healthcare Research.
- Fazel, S., Langstrom, N., Hjern, A., Grann, M., & Lichtenstein, P. (2009). Schizophrenia, substance abuse, and violent crime. *JAMA*, 301(19), 2016-2023.
- Fazel, S., Lichtenstein, P., Grann, M., Goodwin, G. M., & Långström, N. (2010). Bipolar disorder and violent crime: New evidence from population-based longitudinal studies and systematic review. *Archives of General Psychiatry*, 67(9), 931-938.
- Feng, W., Zhou, W., Butler, J. S., Booth, B. M., & French, M. T. (2001). The impact of problem drinking on employment. *Health Economics*, 10(6), 509-521.
- Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2008). Exposure to childhood sexual and physical abuse and adjustment in early adulthood. *Child Abuse and Neglect*, 32(6), 607-619.
- Fergusson, D. M., & Horwood, L. J. (1997). Early onset cannabis use and psychosocial adjustment in young adults. *Addiction*, 92(3), 279-296.
- Fergusson, D. M., & Horwood, L. J. (2000). Alcohol abuse and crime: A fixed-effects regression analysis. *Addiction*, 95(10), 1525-1536.
- Fergusson, D. M., John, H. L., & Ridder, E. M. (2005). Show me the child at seven: The consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry*, 46(8), 837-849.
- Fergusson, D. M., & Lynskey, M. T. (1997). Physical punishment/maltreatment during childhood and adjustment in young adulthood. *Child Abuse & Neglect*, 21(7), 617-630.
- Fergusson, D. M., & Lynskey, M. T. (1998). Conduct problems in childhood and psychosocial outcomes in young adulthood: A prospective study. *Journal of Emotional and Behavioral Disorders*, 6(1), 2-18.
- Fletcher, J. M. (2009). Childhood mistreatment and adolescent and young adult depression. *Social Science and Medicine*, 68(5), 799-806.
- Fletcher, J. M. (2010). Adolescent depression and educational attainment: Results using sibling fixed effects. *Health Economics*, 19(7), 855-871.
- Fletcher, J., & Wolfe, B. (2008). Child mental health and human capital accumulation: The case of ADHD revisited. *Journal of Health Economics*, 27(3), 794-800.
- Fletcher, J. M., & Wolfe, B. L. (2009a). Long-term consequences of childhood ADHD on criminal activities. *The Journal of Mental Health Policy and Economics*, 12(3), 119-138.
- Fletcher, J. M., & Wolfe, B. L. (2009b). Education and labor market consequences of teenage childbearing: Evidence using the timing of pregnancy outcomes and community fixed effects. *Journal of Human Resources*, 44(2), 303-325.
- Forbes, M., Barker, A., & Turner, S. (2010, March). *The effects of education and health on wages and productivity* (Productivity Commission Staff Working Paper). Melbourne, Victoria, Australia: Productivity Commission.
- Francesconi, M. (2008). Adult outcomes for children of teenage mothers. *The Scandinavian Journal of Economics*, 110(1), 93-117.
- Frank, R., & Gertler, P. (1991). An assessment measurement error bias for estimating the effect of mental distress on income. *The Journal of Human Resources*, 26(1), 154-164.
- French, M. T., Roebuck, M. C., & Alexandre, P. K. (2001). Illicit drug use, employment, and labor force participation. *Southern Economic Journal*, 68(2), 349-368.
- French, M. T., Zarkin, G. A. (1998). Mental health, absenteeism and earnings at a large manufacturing worksite. *The Journal of Mental Health Policy and Economics*, 1(4), 161-172.
- French, M. T., Maclean, J. C., Sindelar, J. L., & Fang, H. (2011). The morning after: Alcohol misuse and employment problems. *Applied Economics*, 43(21), 2705-2720.
- Fritjers, P., Johnston, D. W., & Shields, M. A. (2010, April). *Mental health and labour market participation: Evidence from IV panel data models* (IZA Discussion Paper No. 4883). Bonn, Germany: Institute for the Study of Labor.
- Galera, C., Melchior, M., Chastang, J.-F., Bouvard, M.-P., & Fombonne, E. (2009). Childhood and adolescent hyperactivity-inattention symptoms and academic achievement 8 years later: The GAZEL Youth study. *Psychological Medicine*, 39(11), 1895-1906.
- Gibb, S. J., Fergusson, D. M., & Horwood, L. J. (2010). Burden of psychiatric disorder in young adulthood and life outcomes at age 30. *The British Journal of Psychiatry*, 197(2), 122-127.

Studies used in the links meta-analyses

- Goerge, R. M., Harden, A., & Lee, B. J. (2008). Consequences of teen childbearing for child abuse, neglect, and foster care placement. In S. D. Hoffman & R. A. Maynard (Eds.), *Kids having kids: Economic costs & social consequences of teen pregnancy* (2nd ed., 257-288). Washington, DC: Urban Institute Press.
- Hirschfield, P. (2009). Another way out: The impact of juvenile arrests on high school dropout. *Sociology of Education*, 82(4), 368-393.
- Hjalmarsson, R. (2008). Criminal justice involvement and high school completion. *Journal of Urban Economics*, 63(2), 613-630.
- Hoffman, S. D. (2008). Consequences of teen childbearing for mothers (part II: Updated estimates of the consequences of teen childbearing for mothers). In S. D. Hoffman & R. A. Maynard (Eds.), *Kids having kids: Economic costs & social consequences of teen pregnancy* (2nd ed., 74-118). Washington, DC: Urban Institute Press.
- Hoffman, S. D., & Scher, L. S. (2008). Consequences of teen childbearing for the life chances of children, 1979-2002. In S. D. Hoffman & R. A. Maynard (Eds.), *Kids having kids: Economic costs & social consequences of teen pregnancy* (2nd ed., 342-357). Washington, DC: Urban Institute Press.
- Horwitz, A. V., Widom, C. S., McLaughlin, J., & White, H. R. (2001). The impact of childhood abuse and neglect on adult mental health: A prospective study. *Journal of Health and Social Behavior*, 42(2), 184-201.
- Horwood, L. J., Fergusson, D. M., Hayatbakhsh, M. R., Najman, J. M., Coffey, C., Patton, G. C., . . . Hutchinson, D. M. (2010). Cannabis use and educational achievement: Findings from three Australasian cohort studies. *Drug and Alcohol Dependence*, 110(3), 247-253.
- Jofre-Bonet, M., Busch, S. H., Falba, T. A., & Sindelar, J. L. (2005). Poor mental health and smoking: interactive impact on wages. *The Journal of Mental Health Policy and Economics*, 8(4), 193-203.
- Johansson, E., Alho, H., Kiiskinen, U., & Poikolainen, K. (2007). The association of alcohol dependency with employment probability: Evidence from the population survey 'Health 2000 in Finland'. *Health Economics*, 16(7), 739-754.
- Jones, A. S., & Richmond, D. W. (2006). Causal effects of alcoholism on earnings: Estimates from the NLSY. *Health Economics*, 15(8), 849-871.
- Jonson-Reid, M., Drake, B., Kim, J., Porterfield, S., & Han, L. (2004). A prospective analysis of the relationship between reported child maltreatment and special education eligibility among poor children. *Child Maltreatment*, 9(4), 382-394.
- Keng, S. H., & Huffman, W. E. (2010). Binge drinking and labor market success: A longitudinal study on young people. *Journal of Population Economics*, 23(1), 303-322.
- Kenkel, D. S., & Ribar, D. C. (1994). Alcohol consumption and young adults' socioeconomic status. *Brookings Papers on Economic Activity: Microeconomics*, 1994, 119-161.
- Kirk, D. S., & Sampson, R. J. (2009). *Cumulative disadvantage in the adolescent life-course: The case of juvenile arrest and later educational attainment* (Draft). Paper prepared for Brookings Institution, Project on Social Inequality and Educational Disadvantage.
- Lansford, J. E., Dodge, K. A., Pettit, G. S., Bates, J. E., Crozier, J., & Kaplow, J. (2002). A 12-year prospective study of the long-term effects of early child physical maltreatment on psychological, behavioral, and academic problems in adolescence. *Archives of Pediatrics and Adolescent Medicine*, 156(8), 824-830.
- Lansford, J. E., Miller-Johnson, S., Berlin, L. J., Dodge, K. A., Bates, J. E., & Pettit, G. S. (2007). Early physical abuse and later violent delinquency: A prospective longitudinal study. *Child Maltreatment*, 12(3), 233-245.
- Lemmon, J. (1999). How child maltreatment affects dimensions of juvenile delinquency in a cohort of low-income urban youths. *Justice Quarterly*, 16(2), 357-376.
- Levine, J. A., Emery, C. R., & Pollack, H. (2007). The well-being of children born to teen mothers. *Journal of Marriage and Family*, 69(1), 105-122.
- Levitt, S., & Lochner, L. (2001). The determinants of juvenile crime. In J. Gruber (Ed.), *Risky behavior among youths: An economic analysis* (pp. 327-374). Chicago: University of Chicago Press.
- Lipsey, M. W., Wilson, D. B., Cohen, M. A., & Derzon, J. H. (1997). Is there a causal relationship between alcohol and violence? A synthesis of evidence. In M. Galanter (Series Ed.), *Recent Developments in Alcoholism: Vol 13. Alcoholism and violence* (pp. 245-282). New York: Plenum Press.
- Lochner, L., Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review*, 94(1), 155-189.
- MacDonald, Z., & Shields, M. A. (2004). Does problem drinking affect employment? Evidence from England. *Health Economics*, 13(2), 139-155.
- Machin, S., Marie, O., & Vujic, S. (2011). The crime reducing effect of education. *Economic Journal*, 121(552), 463-484.
- Manlove, J. S., Terry-Humen, E., Mincieli, L. A., & Moore, K. A. (2008). Outcomes for children of teen mothers from kindergarten through adolescence. In S. D. Hoffman & R. A. Maynard (Eds.), *Kids having kids: Economic costs & social consequences of teen pregnancy* (2nd ed., 161-220). Washington, DC: Urban Institute Press.
- Marcotte, D. E., & Wilcox-Gök, V. (2003). Estimating earnings losses due to mental illness: A quantile regression approach. *The Journal of Mental Health Policy and Economics*, 6(3), 123-134.
- Maxfield, M. G., & Widom, C. S. (1996). The cycle of violence: Revisited 6 years later. *Archives of Pediatrics and Adolescent Medicine*, 150(4), 390-395.
- McCaffrey, D. F., Pacula, R. L., Han, B., & Ellickson, P. (2010). Marijuana use and high school dropout: The influence of unobservables. *Health Economics*, 19(11), 1281-1299.
- McGloin, J. M., & Widom, C. S. (2001). Resilience among abused and neglected children grown up. *Development and Psychopathology*, 13(4), 1021-1038.
- McLeod, J. D., & Kaiser, K. (2004). Childhood emotional and behavioral problems and educational attainment. *American Sociological Review*, 69(5), 636-658.
- Mensch, B. S., & Kandel, D. B. (1988). Dropping out of high school and drug involvement. *Sociology of Education*, 61(2), 95-113.
- Mersky, J. P., & Reynolds, A. J. (2007). Child maltreatment and violent delinquency: Disentangling main effects and subgroup effects. *Child Maltreatment*, 12(3), 246-258.
- Mersky, J. P., & Topitzes, J. (2010). Comparing early adult outcomes of maltreated and non-maltreated children: A prospective longitudinal investigation. *Children and Youth Services Review*, 32(8), 1086-1096.
- Moore, K. A., Morrison, D. R., & Greene, A. D. (1997). Effects on the children born to adolescent mothers. In R. A. Maynard (Ed.), *Kids having kids: Economic costs and social consequences of teen pregnancy* (pp. 145-180). Washington, DC: Urban Institute Press.
- Mullahy, J., & Sindelar, J. L. (1991). Gender differences in labor market effects of alcoholism. *The American Economic Review*, 81(2), 161-165.
- Murray, J., Irving, B., Farrington, D. P., Colman, I., & Bloxson, C. A. J. (2010). Very early predictors of conduct problems and crime: Results from a national cohort study. *Journal of Child Psychology and Psychiatry*, 51(11), 1198-1207.

Studies used in the links meta-analyses

- Needham, B. L. (2009). Adolescent depressive symptomatology and young adult educational attainment: An examination of gender differences. *Journal of Adolescent Health, 45*(2), 179-186.
- Noll, J. G., Trickett, P. K., & Putnam, F. W. (2003). A prospective investigation of the impact of childhood sexual abuse on the development of sexuality. *Journal of Consulting and Clinical Psychology, 71*(3), 575-586.
- Ojeda, V. D., Frank, R. G., McGuire, T. G., & Gilmer, T. P. (2010). Mental illness, nativity, gender and labor supply. *Health Economics, 19*(4), 396-421.
- Ou, S.-R., & Reynolds, A. J. (2010). Grade retention, postsecondary education, and public aid receipt. *Educational Evaluation and Policy Analysis, 32*(1), 118-139.
- Peters, B. L. (2004). Is there a wage bonus from drinking? Unobserved heterogeneity examined. *Applied Economics, 36*(20), 2299-2315.
- Pogarsky, G., Lizotte, A. J., & Thornberry, T. P. (2003). The delinquency of children born to young mothers: Results from the Rochester Youth Development Study. *Criminology, 41*(4), 1249-1286.
- Porche, M. V., Fortuna, L. R., Lin, J., & Alegria, M. (2011). Childhood trauma and psychiatric disorders as correlates of school dropout in a national sample of young adults. *Child Development, 82*(3), 982-998.
- Renna, F. (2008). Alcohol abuse, alcoholism, and labor market outcomes: Looking for the missing link. *Industrial & Labor Relations Review, 62*(1), 92-103.
- Ringel, J. S., Ellickson, P. L., & Collins, R. L. (2006). The relationship between high school marijuana use and annual earnings among young adult males. *Contemporary Economic Policy, 24*(1), 52-63.
- Sabates, R. (2008). Educational attainment and juvenile crime. *The British Journal of Criminology, 48*(3), 395-409.
- Saffer, H., & Dave, D. (2005). The effect of alcohol consumption on the earnings of older workers. *Advances in Health Economics and Health Services Research, 16*, 61-90.
- Sangchai, C. (2006). *The causal effect of alcohol consumption on employment status*. Tampa: University of South Florida Scholar Commons, Theses and Dissertations.
- Satterfield, J. H., Faller, K. J., Crinella, F. M., Schell, A. M., Swanson, J. M., & Homer, L. D. (2007). A 30-year prospective follow-up study of hyperactive boys with conduct problems: Adult criminality. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(5), 601-610.
- Savoca, E., & Rosenheck, R. (2000). The civilian labor market experiences of Vietnam-era veterans: The influence of psychiatric disorders. *The Journal of Mental Health Policy and Economics, 3*(4), 199-207.
- Scher, L. S., & Hoffman, S. D. (2008). Consequences of teen childbearing for incarceration among adult children (part II: Updated estimates through 2002). In S. D. Hoffman & R. A. Maynard (Eds.), *Kids having kids: Economic costs & social consequences of teen pregnancy* (2nd ed., 311-321). Washington, DC: Urban Institute Press.
- Scott, K. M., Smith, D. R., & Ellis, P. M. (2010). Prospectively ascertained child maltreatment and its association with DSM-IV mental disorders in young adults. *Archives of General Psychiatry, 67*(7), 712-719.
- Shin, S. H., Edwards, E. M., & Heeren, T. (2009). Child abuse and neglect: Relations to adolescent binge drinking in the national longitudinal study of Adolescent Health (AddHealth) Study. *Addictive Behaviors, 34*(3), 277-280.
- Slade, E. P., & Wissow, L. S. (2007). The influence of childhood maltreatment on adolescents' academic performance. *Economics of Education Review, 26*(5), 604-614.
- Springer, K. W., Sheridan, J., Kuo, D., & Carnes, M. (2007). Long-term physical and mental health consequences of childhood physical abuse: Results from a large population-based sample of men and women. *Child Abuse & Neglect, 31*(5), 517-530.
- Steadman, H. J., Mulvey, E. P., Monahan, J., Robbins, P. C., Appelbaum, P. S., Grisso, T., . . . Silver, E. (1998). Violence by people discharged from acute psychiatric inpatient facilities and by others in the same neighborhoods. *Archives of General Psychiatry, 55*(5), 393-401.
- Stouthamer-Loeber, M., Loeber, R., Homish, D. L., & Wei, E. (2001). Maltreatment of boys and the development of disruptive and delinquent behavior. *Development and Psychopathology, 13*(4), 941-955.
- Stouthamer-Loeber, M., Wei, E. H., Homish, D. L., & Loeber, R. (2002). Which family and demographic factors are related to both maltreatment and persistent serious juvenile delinquency? *Children's Services: Social Policy, Research, and Practice, 5*(4), 261-272.
- Tanner, J., Davies, S., & O'Grady, B. (1999). Whatever happened to yesterday's rebels? Longitudinal effects of youth delinquency on education and employment. *Social Problems, 46*(2), 250-274.
- Terza, J. V. (2002). Alcohol abuse and employment: A second look. *Journal of Applied Econometrics, 17*(4), 393-404.
- Thornberry, T. P., Henry, K. L., Ireland, T. O., & Smith, C. A. (2010). The causal impact of childhood-limited maltreatment and adolescent maltreatment on early adult adjustment. *Journal of Adolescent Health, 46*(4), 359-365.
- Thornberry T. P., Ireland, T. O., & Smith, C. A. (2001). The importance of timing: The varying impact of childhood and adolescent maltreatment on multiple problem outcomes. *Development and Psychopathology, 13*(4), 957-979.
- Tian, H., Robinson, R. L., & Sturm, R. (2005). Labor market, financial, insurance and disability outcomes among near elderly Americans with depression and pain. *The Journal of Mental Health Policy and Economics, 8*(4), 219-228.
- Topitzes, J., Mersky, J. P., & Reynolds, A. J. (2010). Child maltreatment and adult cigarette smoking: A long-term developmental model. *Journal of Pediatric Psychology, 35*(5), 484-498.
- Van Dorn, R., Volavka, J., & Johnson, N. (2012). Mental disorder and violence: Is there a relationship beyond substance use? *Social Psychiatry and Psychiatric Epidemiology, 47*(3), 487-503.
- Van Ours, J. C. (2004). A pint a day raises a man's pay; but smoking blows that gain away. *Journal of Health Economics, 23*(5), 863-886.
- Webbink, D., Martin, N. G., & Visscher, P. M. (2009). Does teenage childbearing reduce investment in human capital? *Journal of Population Economics, 24*(2), 701-730.
- Webbink, D., Koning, P., Vujic, S., & Martin, N. (2008, November). *Why are criminals less educated than non-criminals?: Evidence from a cohort of young Australian twins* (CPB Discussion Paper No. 114). The Hague, The Netherlands: CPB Netherlands Bureau for Economic Policy Analysis.
- Webbink, D., Vujic, S., Koning, P., & Martin, N. G. (2011). The effect of childhood conduct disorder on human capital. *Health Economics*. Advance online publication. DOI: 10.1002/hec.1767
- Widom, C. S., DuMont, K., & Czaja, S. J. (2007). A prospective investigation of major depressive disorder and comorbidity in abused and neglected children grown up. *Archives of General Psychiatry, 64*(1), 49-56.
- Widom, C. S., & Kuhns, J. B. (1996). Childhood victimization and subsequent risk for promiscuity, prostitution, and teenage pregnancy: A prospective study. *American Journal of Public Health, 86*(11), 1607-1612.
- Yamada, T., Kendix, M., & Yamada, T. (1996). The impact of alcohol consumption and marijuana use on high school graduation. *Health Economics, 5*(1), 77-92.

Studies used in the links meta-analyses

- Zarkin, G. A., French, M. T., Mroz, T., & Bray, J. W. (1998). Alcohol use and wages: New results from the National Household Survey on Drug Abuse. *Journal of Health Economics*, 17(1), 53-68.
- Zhang, X., Zhao, X., & Harris, A. (2009). Chronic diseases and labour force participation in Australia. *Journal of Health Economics*, 28(1), 91-108.
- Zuvekas, S, Cooper, P. F., & Buchmueller, T.C. (2005, April). *Health behaviors and labor market status: The impact of substance abuse* (Working Paper No. 05013). Rockville, MD: Agency for Healthcare Research and Quality.

Appendix F: Methods to Access Risk and Uncertainty

The model described thus far in this Appendix produces single-point estimates of benefits and costs of different policy and program options. For example, the model may produce an expected bottom line of \$2.35 of benefits for each dollar of costs for some particular program. A key question, however, is this: how risky is an estimate such as this? If we vary the inputs, how often will costs exceed benefits, rather than the other way around?

The Institute's benefit-cost model includes many inputs and assumptions, and there is significant uncertainty around many of these factors. If they are varied, the model will produce different results. Therefore, it is important to test the model systematically for the riskiness inherent in the single point estimates.

We do this by employing a Monte Carlo simulation method where we run the model hundreds of times, each time varying the inputs randomly after sampling from estimated ranges of uncertainty that surround the key inputs. We then record the results of each run of the model.

When this simulation process is complete, we then compute an expected net present value, an expected benefit-cost ratio, and expected internal rate of return, and a straightforward measure of investment risk: for any program, what percentage of the time can we expect benefits to exceed costs? That is, after running the model many times, what percentage of the time will the net present value of benefits be greater than zero (or the benefit-cost ratio be greater than one)?

F1. Key Inputs Varied in the Monte Carlo Simulation Analysis

Potentially, all inputs to the Institute's model could be varied. Since this would slow the model down considerably, we concentrate on estimating the risk and uncertainty around a set of key inputs to the model. Each simulation run draws randomly from estimated distributions around the following list of inputs.

Program Effect Sizes. As described in Appendices B and C, the model is driven by the estimated effects of programs and policies on certain outcomes. We estimate these effect sizes meta-analytically, and that process produces a random effects standard error around the effect size. We model the adjusted mean effect size and the unadjusted standard error by sampling from a normal probability density distribution.

Linked Effect Sizes. Appendices C and E also describe how the model uses estimates of the way in which certain outcomes relate to the outcomes that we monetize in the benefit-cost model. These "linked" effect sizes are also estimated with standard errors and we use the adjusted mean effect sizes and the unadjusted standard errors to sample from a normal probability density distribution.

Discount Rates. The user can enter three different rates of discount (low, modal, and high) that are used to evaluate future benefits and costs in present value terms. In a single run of the model, the modal discount rate is used. In simulation mode the discount rate is sampled from a triangular probability density distribution.

The mean or modal values for many other model inputs are varied in a Monte Carlo run, including:

- Program costs—triangular distribution.
- Crime victimization costs—triangular distribution.
- Criminal justice system costs—triangular distribution.
- Criminal victimizations per conviction—triangular distribution.
- Value of a statistical life—triangular distribution.
- Deadweight cost of taxation—triangular distribution.
- Labor market earnings from reduction in alcohol disorders—lognormal distribution.
- Labor market earnings from reduction in regular tobacco smoking—lognormal distribution.
- Labor market earnings from reduction in cannabis disorders—lognormal distribution.
- Labor market earnings from reduction in non-cannabis illicit drug disorders—lognormal distribution.
- Expected hospital costs per alcohol, illicit drug, or regular smoking event—triangular distribution.
- Expected emergency department costs per alcohol, illicit drug, or regular smoking event—triangular distribution.
- Expected public treatment costs per alcohol, illicit drug, or regular smoking event—triangular distribution.
- Labor market earnings from one standard deviation increase in test scores—normal distribution.
- Labor market earnings from an extra year of education—normal distribution.
- Causal link between high school graduation and labor market earnings—triangular distribution.

F2. Computational Procedures to Carry Out the Simulation

Since the benefit-cost model is housed in Microsoft Excel® and uses Visual Basic for Applications® (VBA) to carry out computations, the simulation is also implemented within VBA using Excel's various statistical functions. First, a random number between zero and one is generated with Excel's *RANDBETWEEN* function with this procedure:

$$RandomDraw = RANDBETWEEN(1,999)/1000$$

Next, the distribution for the particular model input is sampled. For the normal distribution, Excel's normal distribution inverse function, *NORMINV*, is used to generate a draw for any outcome that is set to sample from a normal distribution. For example, an effect size for each run *r* in a simulation is generated with this procedure:

$$EffectSize_r = NORMINV(RandomDraw, EffectSizeMean, EffectSizeStandardError)$$

Other types of probability distributions are computed in a similar fashion.

Excel does not have a native probability function for a triangular distribution. Therefore, the following procedure is used to generate a draw from three triangular parameters supplied by the user. An example would be for the discount rate, *DISRATE*, variable included in simulation runs. VBA implements the following code to randomly draw a discount rate from a triangular distribution given Min, Mode, and Max parameters.

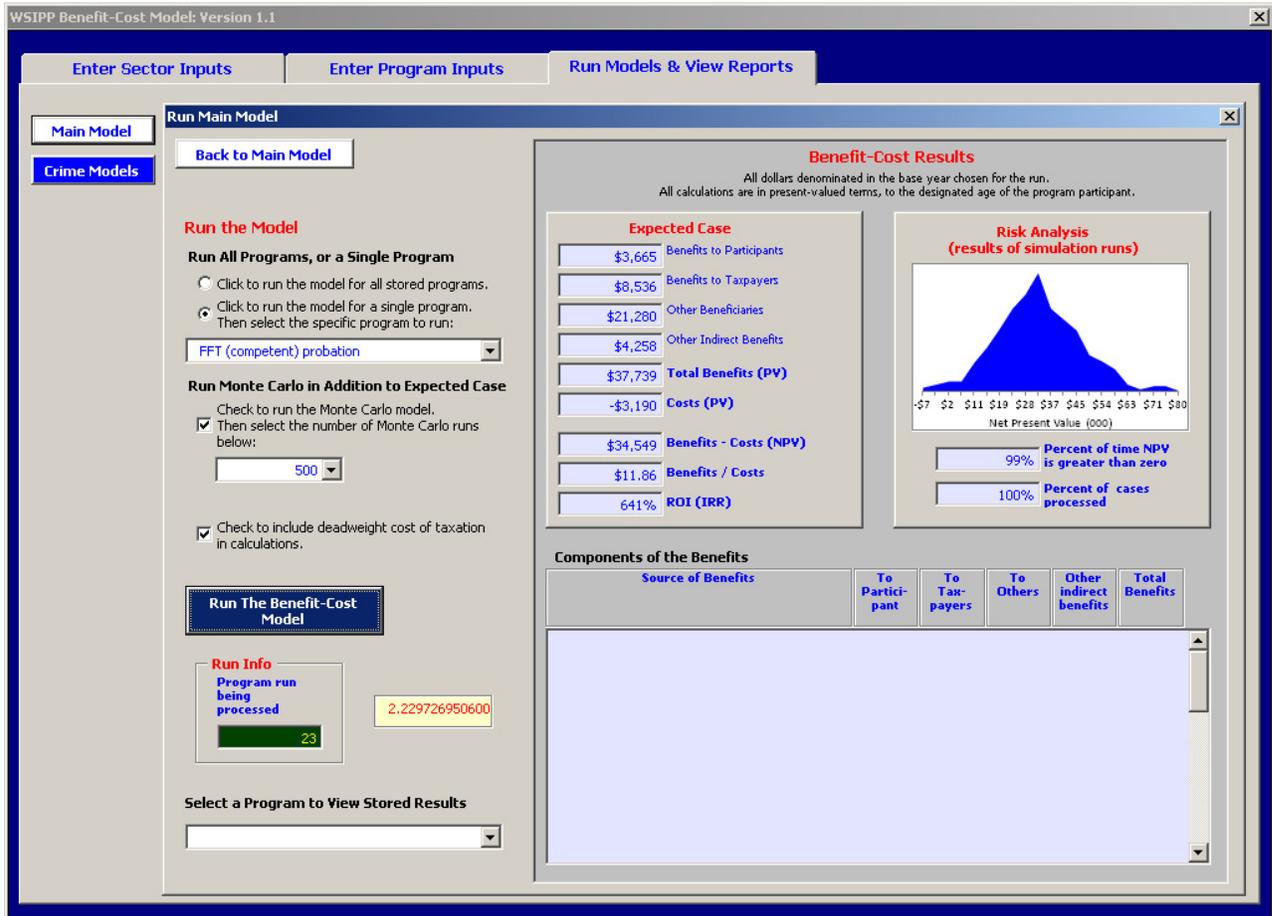
$$\text{If } RandomDraw < \frac{(Mode - Min)}{(Max - Min)} \text{ then } DISRATE = Min + \sqrt{RandomDraw \times (Mode - Min) \times (Max - Min)}$$

$$\text{If } RandomDraw \geq \frac{(Mode - Min)}{(Max - Min)} \text{ then } DISRATE = Max - \sqrt{(1 - RandomDraw) \times (Max - Mode) \times (Max - Min)}$$

Example

When the model is run for a program, and the user selects to run the model in Monte Carlo mode, the user selects the number of runs to compute. Exhibit F1.a is a screen shot of a program, Functional Family Therapy that was run in Monte Carlo mode. The user clicked Monte Carlo on and selected 500 runs. The Monte Carlo results are shown. The chart, for example, displays the distribution of the 500 cases for the Net Present Value for the Program.

Exhibit F1.a

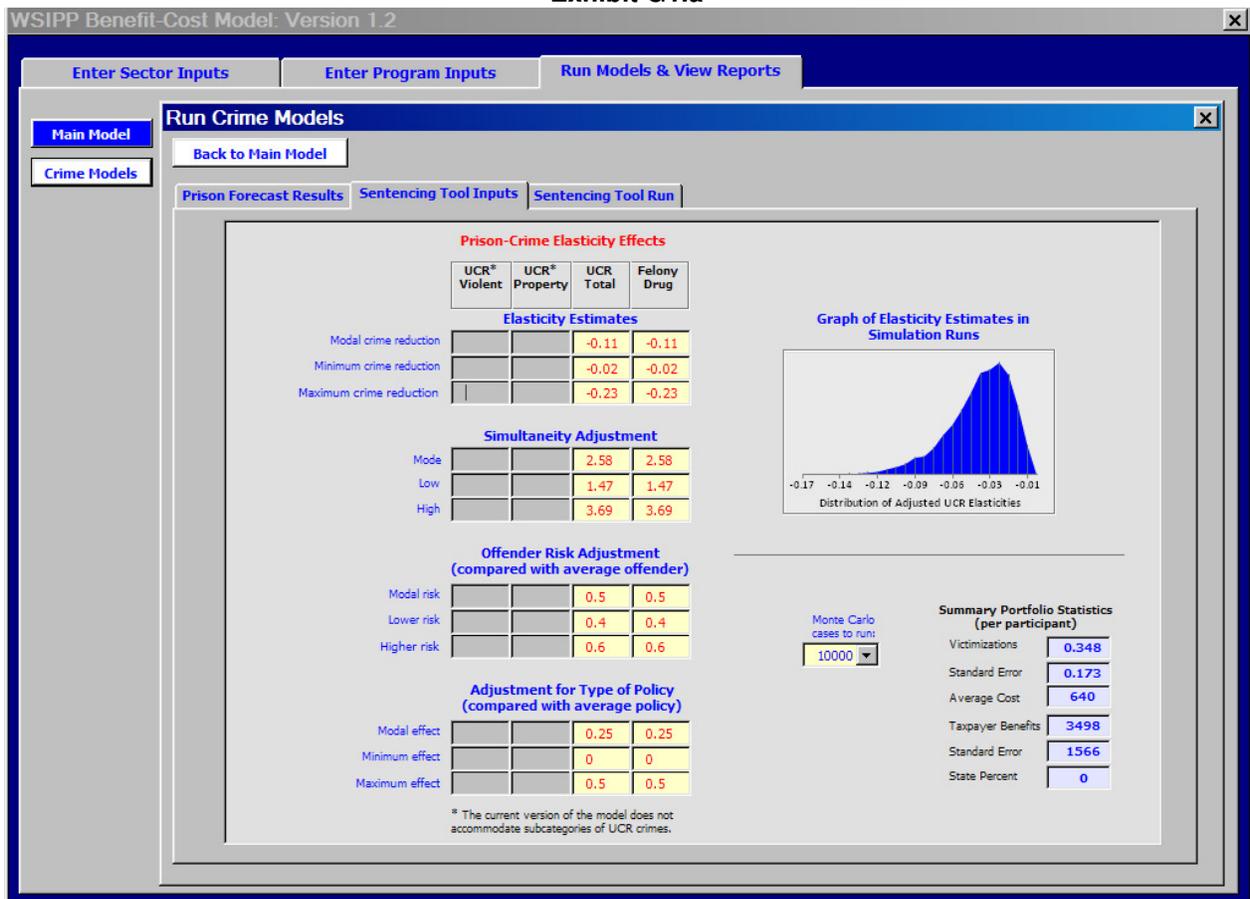


Appendix G: The WSIPP Sentencing and Programming Portfolio Tool

The Washington State Institute for Public Policy has constructed an analytical tool for the Washington legislature to help identify evidence-based sentencing and programming policy options to reduce crime and taxpayer criminal justice costs.¹⁴⁸ The goal of the tool is to help users analyze the net effects of two fundamental types of criminal justice policies available to states: sentencing-related policies and “programming” policies. As we indicate, there is evidence that both types of policies can reduce crime and both, of course, cost taxpayers money. The tool is designed to examine how changes in the mix of these two resources can affect, at the state level, the following: (a) the number victimizations from crime, and (b) taxpayer costs. The sentencing tool resides within the Institute’s benefit-cost model.

Prison-Crime Elasticity Estimates. This screen displays the inputs that collectively estimate the degree to which prison average daily population is expected to affect crime levels in a state. The user enters the estimates of prison-crime elasticities and related adjustments. As we later note, there is uncertainty in all of these parameters. As a result, the tool employs user-supplied low, modal, and high estimates of each of the key parameters. These are then combined in a Monte Carlo simulation to calculate the probability distribution of the expected effectiveness of prison average daily population on crime levels. This screen allows the user to enter these values and then see the resulting distribution of elasticities the parameters produce. The specific values of the input parameters chosen by WSIPP are discussed below.

Exhibit G1.a



¹⁴⁸ S. Aos & E. Drake (2010). *WSIPP's Benefit-Cost Tool for States: Examining Policy Options in Sentencing and Corrections*. Olympia: Washington State Institute for Public Policy, Document No. 10-08-1201.

Evidence-Based Program Portfolio. The second screen displays the pre-loaded information on the effectiveness of various types of correctional programming on crime. We have chosen a selection of programs, for both adult offenders and juvenile offenders, for which we have found credible evidence that they can reduce crime. The screen shows the key summary statistics for each program available to be included in a portfolio. When running the application, the user selects a portfolio and also selects the portion of savings to be applied to fund the chosen portfolio.

Net Impacts on Crime Victimitizations and Taxpayer Costs. The bottom section of the screen displays key statistics summarizing the impacts of the sentencing and programming policy choices.

Exhibit G1.b

WSIPP Benefit-Cost Model: Version 1.2

Enter Sector Inputs | Enter Program Inputs | Run Models & View Reports

Main Model | Crime Models

Run Crime Models

Back to Main Model

Prison Forecast Results | Sentencing Tool Inputs | Sentencing Tool Run

Evidence-Based Program Portfolio Selection

Enter % of fiscal savings from prison ADP change used to purchase program slots:

| Program | Program Cost | Victimizations Avoided | | | Taxpayer Benefits | | | Victim Benefits | Percent of Portfolio | Number of Slots Funded |
|--|--------------|------------------------|-------|------------|-------------------|---------|------------|-----------------|----------------------|------------------------|
| | | Mean | StdEr | Per \$1000 | Mean | StdEr | State Pct. | | | |
| Vocational Education in Prison | \$1,536 | 0.46 | 0.09 | 0.30 | \$4,906 | \$703 | 0% | \$12,569 | 25% | 174 |
| Intensive Supervision with Treatment (High) | \$4,217 | 1.00 | 0.55 | 0.24 | \$8,147 | \$4,137 | 0% | \$21,214 | 0 | 0 |
| Correctional Education in Prison (basic or pos | \$1,102 | 0.49 | 0.10 | 0.45 | \$4,963 | \$1,148 | 0% | \$13,267 | 0 | 0 |
| Cognitive Behavioral Programs in Prison | \$217 | 0.26 | 0.11 | 1.19 | \$2,679 | \$1,103 | 0% | \$6,971 | 25% | 1232 |
| Correctional Industries in Prison | \$1,387 | 0.15 | 0.05 | 0.11 | \$1,546 | \$487 | 0% | \$4,197 | 0 | 0 |
| Drug Treatment in Prison | \$3,893 | 0.35 | 0.06 | 0.09 | \$3,459 | \$701 | 0% | \$9,460 | 0 | 0 |
| Drug Treatment in Community | \$2,102 | 0.30 | 0.08 | 0.14 | \$3,671 | \$919 | 0% | \$9,966 | 0 | 0 |
| Drug Courts (adults) | \$4,095 | 0.20 | 0.04 | 0.05 | \$2,511 | \$275 | 0% | \$8,022 | 0 | 0 |
| Employment Training/Job Assistance in Com | \$132 | 0.09 | 0.03 | 0.67 | \$970 | \$367 | 0% | \$2,988 | 0 | 0 |
| Multidimensional Treatment Foster Care | \$7,730 | 0.73 | 0.73 | 0.09 | \$7,747 | \$5,929 | 0% | \$23,902 | 0 | 0 |
| Family Integrated Transitions (IRA) | \$10,993 | 0.54 | 0.24 | 0.05 | \$5,681 | \$2,066 | 0% | \$17,553 | 0 | 0 |
| Coordination of Services | \$386 | 0.07 | 0.11 | 0.17 | \$786 | \$1,216 | 0% | \$2,247 | 0 | 0 |

Enter portfolio percent for selected program:

Annual prison ADP impact:

Percent change from current ADP:

Check to include capital costs

Check to include local cjs dollars

Run the Model

Impact on Taxpayer Costs

Direct (near-term) Fiscal Impact

Change in prison costs from the ADP policy change:

Change in evidence-based program portfolio costs:

Net change in direct (near-term) costs:

Indirect (long-term) Fiscal Impact

Present value of additional criminal justice costs from the ADP policy change:

Present value of criminal justice costs from the evidence-based portfolio:

Total change in taxpayer costs:

Impact on Victimitizations

Risk Analysis (results of simulation runs)

Average Change in Victimitizations

Change from sentencing policy:

Change from program portfolio:

Net impact on victimizations:

Percent change in current crime rate:

Percent of cases when net victimizations are lower:

Negative numbers = fewer victimizations

The WSIPP sentencing and programming portfolio tool implements a five-step computational process to estimate the following:

1. Effect of prison ADP changes on crime levels
2. Fiscal effect stimulated by these crime changes
3. Benefits and costs of evidence-based crime programs are estimated focusing on their ability to affect crime outcomes and related taxpayer net savings
4. Results (on victimizations and taxpayer costs) of an overall portfolio of sentencing and programming resources are tallied
5. Sensitivity by simulating uncertainty in the analysis using a Monte Carlo approach.

G1. Estimating the Crime Effects of Sentencing-Related Policies (Step 1 of 5)

The analytical task for Step 1 is to estimate the change in the number of crimes that a state can expect to see if incarceration rates are changed. The public policies considered with WSIPP's tool are those sentencing-related policies that affect the overall incarceration rate in a state. The incarceration rate is simply the number of offenders in prison at any point in time divided by a statewide population total (such as total population, or all adults over the age of 18).

Public policies that affect the incarceration rate can be sentencing laws that determine or affect the probability that convicted offenders go to prison, as well as policies that govern the length of sentences. They can also be laws that allow the executive or judicial branch discretion to shorten or, in some cases, lengthen sentences. Since the current population in prison is determined by how many people go to prison and for how long, then all of these sentencing-related public policies collectively affect a state's overall incarceration rate at any point in time. There are other public policies that affect the incarceration rate; for example, early childhood education has been shown to reduce crime and, as a result, can be expected to affect future incarceration rates.¹⁴⁹ For Step 1, however, the focus is on those sentencing-related public policies that directly affect the incarceration rate.

The types of sentencing policy changes may affect the sentences of particular types of offenders, or the policies may affect all offenders. For example, sentencing policy changes may be specifically targeted for drug offenders, or they may be focused on the early release of lower-risk offenders, or they could be general changes that apply to all of those sentenced to prison regardless of offense, criminal history, or risk level. In terms of the expected effect on crime, these are important considerations. We note below that the user can adjust inputs to the tool to better approximate differential sentencing policies.

There is a fairly large amount of research literature on the effect of incarceration rates on crime.¹⁵⁰ Many of the studies addressing this relationship in the United States construct models using state-level data over a number of years to estimate the parameters of an equation of this general form:

$$(G1.1) \quad C_{sy} = a + b(ADP_{sy}) + c(X_{sy}) + e$$

In this typical type of model, crime C in state s and year y is estimated to be a function of a state's overall average daily prison population, ADP , a vector of control variables, X , and an error term, e . Crime is most frequently measured with data from the FBI's Uniform Crime Reporting (UCR) system. The variables are usually divided by population so that they are expressed as rates. The models are almost always estimated with a log-log functional form, at least for the dependent and the main policy variables. Several authors have also observed that the state-level time series data often used to estimate equation (G1.1) are likely have unit roots.¹⁵¹ Thus, to help avoid estimating spurious relationships, some authors estimate equation (G1.1) in first-differences since the time series typically do not exhibit unit roots after differencing once. As noted below, there is also considerable concern in the research literature on the econometric implication of the possible simultaneous relationship between the variables of interest in equation (G1.1): crime may be a function of ADP , but ADP may also be a function of crime. This simultaneous relationship can cause statistically biased estimates if not dealt with.

Marginal effects from this generic log-log crime model are then obtained with:

¹⁴⁹ S. Aos, R. Lieb, J. Mayfield, M. Miller, A. Penucci (2004). *Benefits and Costs of Prevention and Early Intervention Programs for Youth*, Olympia: Washington State Institute for Public Policy, Document No. 04-07-3901.

¹⁵⁰ See Marvell's analysis of 35 studies in: T. B. Marvell (2010). Prison Population and Crime. *Handbook on the Economics of Crime*, B. L. Benson & P. R. Zimmerman (Eds.). Cheltenham, UK: Edward Elgar Publishing.

¹⁵¹ See, for example, Marvell, 2010. See also, W. Spelman (2008). Specifying the relationship between crime and prisons, *Journal of Quantitative Criminology*, 24, 149-178.

$$(G1.2) \Delta C = E \times \frac{(UCR)}{ADP \cdot RRate},$$

In equation (G1.2), the change in crime is estimated with E , the crime-prison elasticity obtained from coefficient b of the typical log-log estimation of equation (G1.1); UCR , the reported crime rate (explained below); ADP , the incarceration rate (explained below), and $RRate$, the reporting rate to police by crime victims (also explained below). The marginal effects are sometimes calculated either at the mean values for ADP - UCR - $RRates$ or, more to the point for policy purposes, at the most recent values for ADP - UCR - $RRates$. The log-log estimation of the constant elasticity E implies diminishing returns when E is less than one and incarceration rates are raised. Similarly, an elasticity less than one coupled with reduced ADP implies increasing returns.

The dependent variable: crime. In the studies estimating these types of equations, crime is most often measured with data from the FBI's UCR. These data count the number of crimes reported to police. Some studies estimate a model of total UCR crime reported to police, while other studies estimate two equations, one for violent crime reported to police, and another for property crime reported to police. Still other studies break the analysis down further and estimate equations for the seven major types of "Part 1" crimes in the UCR data: murder, rape, robbery, aggravated assault, burglary, theft, and motor vehicle theft.

All studies also recognize that not all crime is reported to police and, thus, is not included in the UCR data. Accordingly, most authors, in drawing conclusions from their analyses, use information from the federal National Crime Victimization Survey (NCVS) to obtain estimates about how often crime victims say they report crimes to police. The reporting rate information is then used to adjust the coefficients from the parameters estimated with equation (G1.1) to produce estimates of how the total amount of crime changes as prison population is altered, as depicted in equation (G1.2).

One significant problem with the "Part 1" UCR crime data is that they do not match directly how some states, including Washington, define felony crimes. In Washington, this applies to two types of crimes in particular: felony sex crimes and theft/larceny. The UCR sex offense data only count rapes of females over the age of 12. In addition to this obvious limitation in the UCR data, there are other felony sex crimes (e.g., child molestation), defined by the Revised Code of Washington that are not included in the UCR rape category. Similarly, the UCR data count some types of theft crimes that are below the threshold of felony theft in Washington. As explained in Appendix D.2, we make adjustments for these two crime types, since the policy focus of Washington's sentencing laws is felony crime as defined in Washington.

The policy variable: average daily prison population. In virtually all studies in the research literature, the policy variable of interest is prison average daily population. In most studies, this is measured by counting the total number of inmates at the beginning of a year, or the end of a year, and dividing by a state's population aggregate to obtain an overall incarceration rate. Measuring ADP with the total number of offenders—as opposed to more refined categories of offenders convicted of violent, property, or drug offenses, or defining offenders based on an actuarial risk assessment as high risk, moderate risk, or lower risk offenders—is necessary in cross-state analyses, because total ADP is usually the only information available.

The typical research study only includes a measure of total ADP and, thereby, only measures the average effect on crime of the average offender sentenced to prison. However, the criminal propensities of different types of offenders—for example, sex offenders or property offenders—are quite heterogeneous in terms of the amount and types of crime committed. For example, among offenders in Washington's current average daily population, there are some very high and lower risk offenders.¹⁵² Thus, the "average offender" findings from typical research studies limit the practical policy relevance in crafting specific sentencing policies.

Given trends in sentencing policies in the United States, the "average offender" limitation poses at least three empirical problems. First, the average mix of offenders in prison has changed over time. For example, in Washington State, there were virtually no offenders in prison for drug crimes prior to the mid-1980s. Sentencing laws were changed in the late 1980s and the average proportion of drug offenders in ADP increased substantially. The average risk for reoffense has also exhibited long-term trends. Among offenders released from prison in Washington, there has been a 23 percent increase in offenders' risk level between 1991 and 2005.¹⁵³ Thus, the average crime/ADP coefficient from most regressions may not be aligned with the current mix of offenders in a state's ADP.

The second reason why parameters in models like equation (G1.1) are limited in their ability to inform actual policy choices facing legislatures is that policy decisions to raise or lower ADP are not usually across-the-board or "average" decisions. A legislature will rarely raise sentences for all types of crimes by a uniform amount, nor will a legislature typically lower sentences uniformly for all types of crimes (although this has been done). If a legislature were to uniformly lower lengths of stay, for example, a high risk sex offender would be treated the same as a low risk drug offender. Since this is likely to be

¹⁵² E. Drake, S. Aos, & R. Barnoski (2010). *Washington's Offender Accountability Act: Final Report on Recidivism Outcomes*. Olympia: Washington State Institute for Public Policy, Document No.10-01-1201.

¹⁵³ *Ibid.*

seen as an undesirable policy, if prison ADP is to be adjusted, legislative discussions are more likely to focus on at least some level of policy selectivity in which types of offenders are released early. Much more often, a legislature will adjust sentencing statutes for particular types of crimes, rather than adopt across-the board changes.

The third very important reason why a policy adjustment needs to be made to the average elasticity estimates is that not all policies that affect prison ADP have an equal effect on crime. Durlauf and Nagin (2010) provide a very useful review and analysis of the research literature on the two sentencing factors that determine a state's ADP: the probability of a sentence to prison given a conviction, and the severity of the sentence in terms of length of prison stay.

Each of these sentencing parameters—the certainty of punishment and the severity of punishment—are affected by different sentencing policies. Yet, as Durlauf and Nagin found, the research literature indicates that the two types of policies are likely to have quite different effects on crime. They state:

*The key empirical conclusion of our literature review is that there is relatively little reliable evidence of variation in the severity of punishment having a substantial deterrent effect but that there is relatively strong evidence that variation in the certainty of punishment has a large deterrent effect.*¹⁵⁴

Thus, when using the WSIPP sentencing tool to measure how a change in ADP affects crime, via the estimated elasticities discussed above, it is likely to matter a lot if the policy affecting ADP is based on a change to the certainty or severity of punishment. Using the Durlauf and Nagin results, one would conclude that the mean ADP elasticity for a sentencing policy that affects the certainty of punishment would be higher. Conversely, the mean ADP elasticity for a sentencing policy that affects the length of prison stay would be lower.

While the state of research may not allow a clear delineation of the magnitude of these differential effects, the direction seems clear based on the findings of Durlauf and Nagin. Therefore, the sentencing tool allows the user to enter low, modal, and high ADP policy adjustments to modify the overall elasticities obtained from studies estimating models of the type shown in equation (G1.1).

This means that the coefficients obtained from equations such as (G1.1) above can be thought of as only rough guides for the effectiveness of average sentencing changes. The coefficients obtained from these equations need to be adjusted to better estimate the specific policy choices available to legislatures. Adjustments need to reflect: (a) the wide heterogeneity of criminal propensities among offenders, (b) that legislatures usually adjust sentencing policies differentially for different types of crimes, and (c), that the type of sentencing policy is likely to affect crime differentially depending on whether the policy changes the certainty or severity of punishment. Our approach attempts to account for, even if crudely, some of these necessary policy adjustments.

Simultaneity Considerations. Another major empirical difficulty, observed by many, in providing credible estimates from models like those in equation (G1.1) is related to the likely nature of the relationship between crime levels and prison levels. Crime may be affected by prison, but there is also evidence in many of the studies that the use of prison is affected by crime.¹⁵⁵ This simultaneous relationship, if not accounted for, will probably bias the coefficient in a model like equation (G1.1) downward. If a legislature's willingness to provide prison cells is motivated by changes in crime levels, then the observed relationship between prison and crime can be measuring both prison supply decisions and criminal response to prison levels. Therefore, an observed effect of prison on crime is likely to be muted because some of the observed relationship reflects the use of more prison as a result of crime changes. In the research literature, there have been only a few attempts to measure the magnitude of this simultaneous relationship.¹⁵⁶ Technically, these models require an exogenous source of variation—an instrumental variable or a discontinuity around some arbitrary sentencing cut-off level—that affects the use of prison but is probably otherwise unrelated to the error term in equation (G1.1). These instrumental variables are hard to find so there are many more estimates that do not account for simultaneity than that do. The tool we constructed allows the user to enter different estimates of this ostensibly important effect.

WSIPP's Tool to Estimate Changes in Crime Levels From Changes in Incarceration Rates. The first of five major steps in WSIPP's sentencing tool calculates the expected number of crimes that a state will add or subtract if sentencing policies alter ADP.

An estimate of the total current level of crimes, C_T , in a state is computed with equation (G1.3). Reported UCR crimes for each type of crime c are multiplied by any adjustment to the UCR crime series for each crime type, and summed. This total is divided by the weighted average reporting rate for crimes to compute the estimated total crimes. Since the focus of the

¹⁵⁴ Durlauf & Nagin, 2010.

¹⁵⁵ Marvell, 2010.

¹⁵⁶ S. D. Levitt (1996). The Effect of Prison Population Size on Crime Rates: Evidence From Prison Overcrowding Litigation. *The Quarterly Journal of Economics*, 111, 319-51. W. Spelman (2005). Jobs or Jails? The Crime Drop in Texas. *Journal of Policy Analysis and Management*, 24(1), 133-165.

sentencing ntool is changes to current levels of crime stemming from changes of current levels of incarceration, the current level is computed for the most recent UCR data available.

$$(G1.3) C_T = \frac{\sum_{c=1}^{Ct} (UCR_c \times UCRAAdj_c)}{RRateAdj_T}$$

Since the computation of marginal effects for equation (G1.2) is designed for one unit changes in ADP, and since the tool may be used to estimate the effects of large changes in ADP, the computation of the total marginal crime effect is estimated iteratively, one ADP at a time. Equation (G1.4) implements this iterative process. The tool sums the (absolute value) of a total sentencing change, ΔADP . Equation (G1.4) is similar to equation (G1.2) with additional parameters to explicitly address the issues raised above. For a policy that raises or lowers total prison ADP, the change in total crime, ΔC_T , is calculated with an estimate of the total elasticity, E_T , multiplied by a total simultaneity adjustment, S_T , multiplied by an adjustment, A_T , to account for some level of policy selectivity adopted by a legislative body. S_T is likely to be greater than one while A_T is likely to be less than one if a sentencing-related policy change is selectively applied by a legislature to lower-risk offenders (as opposed to the average risk of the average offender in ADP_T). The marginal effect calculation is then completed by multiplying the product of these three terms by total Crimes at each iteration of the total ADP change. This product is then divided by total ADP at each iteration for the total ADP change. If ADP is increased by a policy change, then ADP increases (+) by one unit for each iteration a ; if ADP is decreased by a policy change, then ADP decreases (-) by one unit for each iteration a .

$$(G1.4) \Delta C_T = \sum_{a=1}^{|\Delta ADP_T|} \frac{(E_T \times S_T \times A_T) \times [C_T - \sum_{i=1}^a (\Delta C_{T(i-1)})]}{(ADP_T \pm a)}$$

For example, for a 100 unit change in ADP, equation (G1.4) is estimated 100 times, each time substituting a one unit difference in ADP and the new level of the UCR variable after the previous delta crime has been computed.

For use in Step 2 of WSIPP's sentencing tool, we are interested in the change in reported crimes, not total crimes. This is given by equation (G1.5).

$$(G1.5) \Delta RC_T = \Delta C_t \times RRateAdj_t$$

Equation (G1.3), which describes total felony crime in a state, can be broken down into crime subcategories. For example, rather than estimating total crime, two equations (with separate inputs) can produce estimates for violent crime and property crime.

$$(G1.6) \Delta C_V = \sum_{a=1}^{|\Delta ADP_T|} \frac{(E_V \times S_V \times A_V) \times [C_V - \sum_{i=1}^a (\Delta C_{V(i-1)})]}{(ADP_T \pm a)}$$

$$(G1.7) \Delta C_P = \sum_{a=1}^{|\Delta ADP_T|} \frac{(E_P \times S_P \times A_P) \times [C_P - \sum_{i=1}^a (\Delta C_{P(i-1)})]}{(ADP_T \pm a)}$$

For the key inputs in equation (G1.3), or equations (G1.6) or (G1.7), WSIPP's tool allows for user-specified uncertainty in the parameters. For example, for the elasticity parameter, the user can specify low, modal, and high parameters. In Monte Carlo simulation, these three parameters are used to randomly draw from a triangular probability density distribution when the equations are estimated. The user can similarly specify low, modal, and high parameters for the simultaneity adjustment and the policy selectivity adjustment.

Equation (G1.4), or equations (G1.6) or (G1.7), contains the following user-specified inputs:

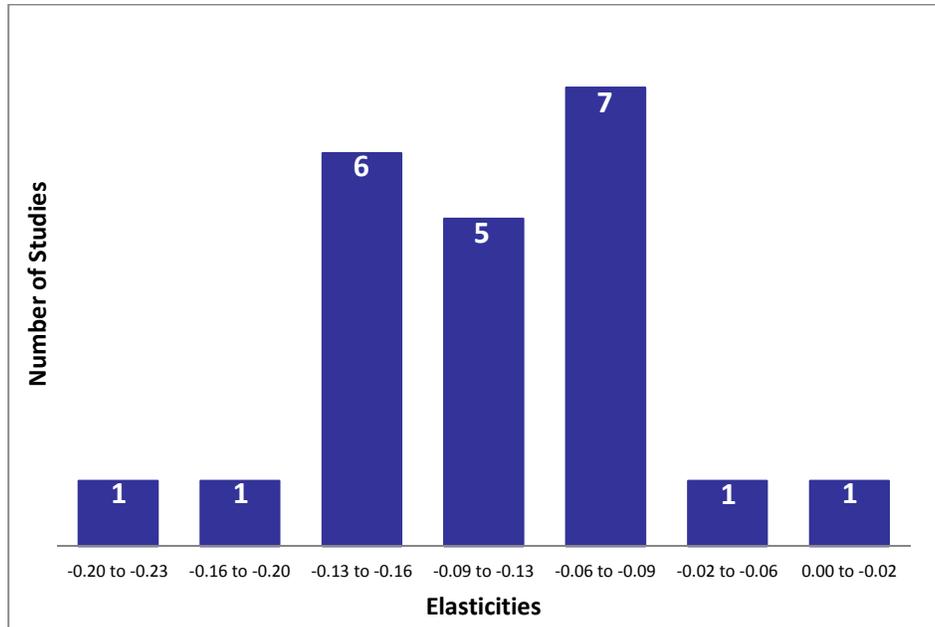
Elasticity Estimates. As noted, there is a fairly sizable research literature that has attempted to estimate the effect of average incarceration on crime. Marvell (2010) recently surveyed the entire literature and summarized the estimates.¹⁵⁷ The chart below summarizes the result of Marvell's review of 22 elasticity estimates of how ADP affects the total crime rate (21 of the estimates are from previously published research studies, one of the 22 estimates was Marvell's new original analysis, as reported in his 2009 review).¹⁵⁸ The findings from the studies shown on this chart did not account for the simultaneity issues discussed above; these studies, which form the existing knowledge base, measure the effect of total prison ADP on total UCR

¹⁵⁷ Marvell, 2010.

¹⁵⁸ We did not include the results of the national level analyses summarized in Marvell; this summary only includes the results of studies using state or local data sets.

crime. The average elasticity estimate for this group of estimates is -0.11 (the median was also -0.11) with a standard deviation of 0.05. Using these results, we entered in the WSIPP sentencing tool the following parameters for a triangular probability distribution: a modal elasticity of -.11 with a minimum crime reduction elasticity of -.02 and maximum crime reduction elasticity of -.23.¹⁵⁹ This range seems to fairly well reflect the current consensus on prison-crime elasticities. It is important to reiterate, however, that these estimates are likely to be biased downward, because they do not reflect the likely simultaneity effect that exists between prison and crime—therefore, we apply a simultaneity adjustment (see below) to these base elasticity estimates.¹⁶⁰

Exhibit G1.c
Summary of the Number of Elasticity Estimates Reported in
the Literature Review of Marvell (2009)



In addition to examining the results of Marvell’s 2009 literature review, we also estimated our own econometric model of the crime-prison relationship for Washington State. Our model estimates an equation similar to equation (G1.1), except that we used county-level data from 1982 to 2008 for Washington’s 39 counties. For this analysis, we ran models with total statewide average daily prison population in Washington (expressed as a rate by dividing by the 18- to 49-year-old population) and total county-level UCR crime (expressed as a rate by dividing by total Washington population). In keeping with the functional forms reviewed in the Marvell study, we estimated a model for the log of county total UCR crime rate and the log of the total statewide ADP prison rate. We also controlled for the log of the local county jail incarceration, the log of county-level police employment, the county unemployment rate, and the log of county real per capita income. The model includes county-level fixed effects. White standard errors are reported. When the log-log model is run in levels, the resulting elasticity is -.07, which is in the range of those observed in the Marvell review.

¹⁵⁹ See Exhibit G1.a.

¹⁶⁰ Spelman, 2008, pg. 168.

Exhibit G1.d

| Dependent Variable: LOG(CR_RATE_TOT) | | | | |
|--|-------------|-----------------------|-------------|-----------|
| Method: Panel Least Squares | | | | |
| Date: 08/30/10 Time: 17:27 | | | | |
| Sample (adjusted): 1983 2008 | | | | |
| Periods included: 26 | | | | |
| Cross-sections included: 39 | | | | |
| Total panel (balanced) observations: 1014 | | | | |
| White diagonal standard errors & covariance (d.f. corrected) | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 14.13445 | 1.414325 | 9.993774 | 0.0000 |
| LOG(ADP_TOTAL1049(-1)) | -0.074764 | 0.081799 | -0.914001 | 0.3609 |
| LOG(ADP_JAIL1849MOD(-1)) | 0.004265 | 0.005725 | 0.745031 | 0.4564 |
| LOG(POLICEEMP1049(-1)) | -0.474798 | 0.096142 | -4.938509 | 0.0000 |
| E_UR | 0.077093 | 0.391337 | 0.197000 | 0.8439 |
| LOG(E_PCIR(-1)) | -0.518642 | 0.148012 | -3.504053 | 0.0005 |
| Effects Specification | | | | |
| Cross-section fixed (dummy variables) | | | | |
| R-squared | 0.767389 | Mean dependent var | | 8.334487 |
| Adjusted R-squared | 0.757078 | S.D. dependent var | | 0.417188 |
| S.E. of regression | 0.205620 | Akaike info criterion | | -0.283151 |
| Sum squared resid | 41.01121 | Schwarz criterion | | -0.069588 |
| Log likelihood | 187.5574 | Hannan-Quinn criter. | | -0.202037 |
| F-statistic | 74.42002 | Durbin-Watson stat | | 0.516363 |
| Prob(F-statistic) | 0.000000 | | | |

We then tested for unit roots in the crime and prison ADP variables. The county total UCR crime rate data did not exhibit unit roots; the Im-Pesaran-Shin panel unit root test produced a p-value of .034, thus rejecting the null hypothesis of a unit root. We then tested the prison ADP variable and we found a strong indication of a unit root with a Im-Pesaran-Shin p-value of 1.0, thus clearly not rejecting the presence of a unit root series. Since the ADP variable is a statewide rate applied to all counties, we also tested the single statewide series from 1982 to 2008 with an Augmented Dickey Fuller test, and we also did not reject a unit root (p-value = .94). With a trend and an intercept, the p-value of the Augmented Dickey Fuller test remained non-significant .74. In first differences, on the other hand, the Augmented Dickey Fuller test had a p-value of .02 and the Im-Pesaran-Shin panel unit root test produced a p-value of .000. Thus a first difference model is indicated based on the ADP variable. The need to use a first difference estimation of the prison-crime relationship is consistent with the recommendation in Marvell (2009) and Spelman (2008). Additional research is warranted with this Washington county level data set to test for possible cointegration.

The results of a first difference specification of the logged variables is shown below. Here the prison elasticity is larger, at -.33. This is in line with some estimates that have been produced when simultaneity is accounted for with instrumental variables. As noted by Spelman (2008), it is possible that by using county level data within a particular state (Washington, in this case), the data may not require accounting for simultaneity. Washington also had an arguably close-to-exogenous change in ADP in the 1980s when it adopted a new form of adult sentencing. When this new system went into effect, the incarceration rate was lowered as a matter of policy; this trend was later reversed by subsequent sentencing policy actions, but for a few years this seemingly exogenous policy change probably allowed a cleaner delineation of the true prison-crime relationship. This may be why the -.33 estimate in the first difference model is close to the estimates obtained by other studies when simultaneity is accounted for.

Exhibit G1.e

| Dependent Variable: CR_RATETOT_FDL | | | | |
|--|-------------|-----------------------|-------------|-----------|
| Method: Panel Least Squares | | | | |
| Date: 08/30/10 Time: 17:21 | | | | |
| Sample (adjusted): 1984 2008 | | | | |
| Periods included: 25 | | | | |
| Cross-sections included: 39 | | | | |
| Total panel (balanced) observations: 975 | | | | |
| White diagonal standard errors & covariance (d.f. corrected) | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | -0.000920 | 0.005504 | -0.167186 | 0.8673 |
| ADP_TOT1049_FDL(-1) | -0.326227 | 0.111908 | -2.915149 | 0.0036 |
| POLICEEMP1049_FDL(-1) | -0.165927 | 0.068704 | -2.415085 | 0.0159 |
| ADP_JAIL1849MOD_FDL(-1) | 0.000695 | 0.006330 | 0.109770 | 0.9126 |
| E_UR-E_UR(-1) | -0.406372 | 0.481598 | -0.843800 | 0.3990 |
| E_PCIR_FDL(-1) | -0.303863 | 0.146536 | -2.073633 | 0.0384 |
| R-squared | 0.029606 | Mean dependent var | | -0.014413 |
| Adjusted R-squared | 0.024599 | S.D. dependent var | | 0.143917 |
| S.E. of regression | 0.142136 | Akaike info criterion | | -1.057935 |
| Sum squared resid | 19.57626 | Schwarz criterion | | -1.027889 |
| Log likelihood | 521.7433 | Hannan-Quinn criter. | | -1.046501 |
| F-statistic | 5.912732 | Durbin-Watson stat | | 2.198526 |
| Prob(F-statistic) | 0.000022 | | | |

The Marvell (2009) review provides the basis for the estimated range of elasticities we include in the WSIPP sentencing tool. The evidence from the regression results for Washington appear to align with Marvell's findings. Additional exploration of the Washington county-level data may provide further insights.

Simultaneity Adjustment. In addition to the studies reported above, Marvell (2009) also reviewed the few studies that have been done using instrumental variables to control for simultaneity bias. The median elasticity estimate for total crime of the three studies that controlled for simultaneity is -.28, compared to the -.11 estimate from the 22 studies that did not control for simultaneity. A simple ratio of these estimates produces a multiple of 2.58; that is, accounting for simultaneity could raise the average elasticity 2.58 times. We use this ratio as a modal, multiplicative simultaneity adjustment in the WSIPP sentencing tool. We bound this rough estimate on the high side with the OLS and IV estimates reported in the well-known study by Levitt (1996). Levitt used a common dataset and method and produced IV elasticities that were 3.69 times higher than his OLS elasticity estimates where he did not control for simultaneity. We use this ratio as our high simultaneity adjustment. For the low simultaneity adjustment, we simply take the difference between 2.58 and 3.69 and subtract it from 2.58. These three multiplicative simultaneity adjustments, 1.47, 2.58, and 3.69, are then used in a triangular probability density distribution to adjust the basic elasticity estimate discussed above. We entered these three parameters in the sentencing tool as shown in Exhibit G1.a.

Average Offender Risk Adjustment. As noted previously, policymakers are not likely to apply uniform sentencing practices for offenders with varying risk levels. For example, if a legislature were to consider lowering prison sentences, it probably would draw some distinction between high risk sex offenders and lower-risk drug offenders. The empirical studies that estimate how a change in ADP affects crime, however, do not consider this type of differential change in ADP; instead, nearly all studies estimate the effect of total ADP on crime. The sentencing tool is designed to allow users to enter parameters, as shown in Exhibit G1.a, that scale the elasticities to better reflect the selective policy adjustments compared with the average offender.

Average Policy Adjustment. As previously discussed, Durlauf and Nagin (2010) found that there is more empirical evidence of a deterrent effect from the certainty of punishment, but no reliable evidence of a deterrent effect for the severity of punishment. Thus, the sentencing tool contains inputs to adjust for the type of policy being considered as shown in Exhibit G1.a. Using the Durlauf and Nagin results, one would conclude that the mean ADP elasticity for a sentencing policy that affects the certainty of punishment would be higher. Conversely, the mean ADP elasticity for a sentencing policy that affects the length of prison stay would be lower.

The Combined Effect of the Range of Elasticity Estimates, Simultaneity Adjustments, and Policy Adjustments. The sentencing tool is designed to be run as a Monte Carlo simulation since there is a great deal of uncertainty in some of the key parameters in the tool. The true effect of prison on crime, as measured by the elasticity, is a major source of uncertainty. The effect can be broken down into the component parts and then combined multiplicatively, and in equation (G1.4). Once the user has entered low, modal, and high parameters for each of the three factors, triangular probability density distributions are created. Then, in Monte Carlo simulation, each of the three parameters is drawn randomly from the triangular distribution and the joint product is computed. The chart shown in Exhibit G1.a displays a distribution of elasticity estimates after a 10,000 case Monte Carlo run, given the parameters.

Uniform Crime Report Adjustments. Not all UCR reported crime categories align with felony conviction data as defined by the Revised Code of Washington. We describe the procedures we use to adjust the UCR data in Appendix D2.3.

Reporting Rate Adjustments. The preceding adjustments to the UCR data also require that we make adjustments to the reporting rates in the National Crime Victimization Survey. We describe these adjustments in Appendix D2.3.

G2. Estimating the Fiscal Effect of Sentencing-Related Policy Changes (Step 2 of 5)

There can be two related fiscal effects that stem from the results of Step 1. If ADP is lowered or raised by a sentencing policy change, then there should be a fiscal effect on state budgets. For example, we have estimated that a one-unit change in ADP can be expected to change state prison operating costs by about \$12,722 (in 2009 dollars) for Washington State. As we explain in Appendix D2, this figure is estimated econometrically and measures the marginal budgetary cost of changes to staffing levels and other operating costs of state prisons; it does not include capital costs. The user can elect to include capital costs in all calculations. If this option is selected (see the input screen), then the user-supplied annualized capital cost is added to this estimate. For Washington, we have estimated that the annual taxpayer capital payment for a prison bed is about \$8,308; the derivation of this estimate is also discussed in Appendix D2. Thus, the first fiscal effect from a change in prison ADP is the budgetary effect on state prison budgets (with or without capital costs included).

The second fiscal effect stemming from a change in prison ADP has to do with other fiscal costs that can be expected to change to the degree that a change in prison ADP affects the crime rate in a state. If sentencing policies reduce incarceration rates, and if this change results in an increase in crime, then some of those new crimes will be processed through the criminal justice system and this will cost taxpayers money. These increased costs will offset the immediate prison fiscal savings from the ADP reduction.

If, on the other hand, changed sentencing policies result in increased incarceration rates, and if this results in a decrease in crime, then other fiscal costs can be expected to be reduced as crime drops. In this case, increased costs associated with the increased ADP will be offset by expected reductions in other fiscal costs as crime goes down.

WSIPP's tool estimates both of these fiscal effects.

Percentage of Reported Crimes That Result in a Prison or Jail Sentence. To estimate net fiscal effects, we begin by estimating the proportion of crimes reported to police that result in prison or jail sentences, via the combined effects of policing and sentencing laws. For total felony crimes T , equations (G2.1) and (G2.2) use information for the most recent year available on the number of prison or jail sentences handed down in a state for the seven felony crime categories. This sum is divided by the UCR data and adjustments, as described in equation (G1.3).

$$(G2.1) \text{PrisonProb}_T = \frac{\sum_{c=1}^{Ct} \text{PrisonSentences}_c}{\sum_{c=1}^{Ct} (\text{UCR}_c \times \text{UCRAdj}_c)}$$

$$(G2.2) \text{JailProb}_T = \frac{\sum_{c=1}^{Ct} \text{JailSentences}_c}{\sum_{c=1}^{Ct} (\text{UCR}_c \times \text{UCRAdj}_c)}$$

The Washington data sources for equations (G2.1) and (G2.2) are described in Appendix D2. Equations (G2.1) and (G2.2) describe the process for the average prison or jail sentence probability for total felony crimes, T . Similar equations, not shown, can be calculated for the violent crime or property crime subcategories.

Average Length of Stay in Prison or Jail. The calculation of fiscal effects for total felony crime T also uses an estimate of the average length of stay for offenders sentenced to prison or jail, by type of felony crime. The data source for Washington State for the variables in (G2.3) and (G2.4) are described in Appendix D2.

$$(G2.3) \text{PrisonLOS}_T = \frac{\sum_{c=1}^{Ct} (\text{PrisonLOS}_c \times \text{PctTimeServed}_c \times \text{PrisonSentences}_c)}{\sum_{c=1}^{Ct} (\text{PrisonSentences}_c)}$$

$$(G2.4) \text{ JailLOS}_T = \frac{\sum_{c=1}^{Ct} (\text{JailLOS}_c \times \text{JailSentences}_c)}{\sum_{c=1}^{Ct} (\text{JailSentences}_c)}$$

Percentage of Prison or Jail Sentences That Receive Community Supervision. In addition to serving institutional time, a person sentenced to prison or jail may also receive a sentence that involves community supervision, which in many states is called parole or probation. The data source for Washington State for the variables in (G2.5) and (G2.6) are described in Appendix D2. The two equations estimate the weighted probabilities for total felony crimes T .

$$(G2.5) \text{ PostPrisonCSProb}_T = \frac{\sum_{c=1}^{Ct} (\text{PostPrisonSentencesCS}_c \times \text{PrisonSentences}_c)}{\sum_{c=1}^{Ct} (\text{PrisonSentences}_c)}$$

$$(G2.6) \text{ PostJailCSProb}_T = \frac{\sum_{c=1}^{Ct} (\text{PostJailSentencesCS}_c \times \text{JailSentences}_c)}{\sum_{c=1}^{Ct} (\text{JailSentences}_c)}$$

Average Length of Stay on Community Supervision. The fiscal-effects model also uses an estimate of the average length of stay for offenders sentenced to prison or jail, by type of felony crime. The data source for Washington State for the variables in (G2.7) and (G2.8) are described in Appendix D2. The two equations estimate the weighted length of stays for total felony crimes T .

$$(G2.7) \text{ PostPrisonCSLOS}_T = \frac{\sum_{c=1}^{Ct} (\text{PostPrisonCSLOS}_c \times \text{PrisonSentences}_c)}{\sum_{c=1}^{Ct} (\text{PostPrisonCSSentences}_c)}$$

$$(G2.8) \text{ PostJailCSLOS}_T = \frac{\sum_{c=1}^{Ct} (\text{PostJailCSLOS}_c \times \text{PostJailCSSentences}_c)}{\sum_{c=1}^{Ct} (\text{PostJailCSSentences}_c)}$$

Change in Prison Costs. The change in the present value of prison costs for the change in all types of felony crime T , given the above equations, is then computed with:

$$(G2.9) \Delta \text{Prison}\$_T = \sum_{y=1}^{\text{PrisonLOS}_T} \frac{\Delta RC_T \times \text{PrisonProb}_T \times \text{Prison}\$}{(1 + \text{dis})^{y-1}}, \text{ if } \text{PrisonLOS}_T < 1, \text{ then } y = \text{PrisonLOS}_T$$

The change in reported crime from the change in prison ADP is computed from equation (G1.5). The variable dis is the discount rate used in the overall benefit-cost analysis, and is entered by the user on the sentencing tool screen. The variable for prison costs, $\text{Prison}\$$, is the calculated figure discussed above and represents changed operating and capital costs (if the user has selected the capital cost inclusion option).

Change in Jail Costs. The change in the present value of jail costs for the change in all types of felony crime T , given the above equations, is then computed with:

$$(G2.10) \Delta \text{Jail}\$_T = \sum_{y=1}^{\text{JailLOS}_T} \frac{\Delta RC_T \times \text{JailProb}_T \times \text{Jail}\$}{(1 + \text{dis})^{y-1}}, \text{ if } \text{JailLOS}_T < 1, \text{ then } y = \text{JailLOS}_T$$

The change in reported crime from the change in prison ADP is computed from equation (G1.5). The variable dis is the discount rate used in the overall benefit-cost analysis, and is entered by the user on the WSIPP tool screen. The variable for jail costs, $\text{Jail}\$$, is the calculated operating costs estimate and capital costs if the user has selected the capital cost inclusion option.

Change in Post-Prison Community Supervision Costs. The change in post-prison supervision costs, given the above inputs is then computed with:

$$(G2.11) \Delta \text{PostPrisonCS}\$_T = \frac{\sum_{y=1}^{\text{PostPrisonCSLOS}_T} \frac{\Delta RC_T \times \text{PrisonProb}_T \times \text{CS}\$}{(1 + \text{dis})^{y-1}}, \text{ if } \text{PrisonLOS}_T < 1, \text{ then } y = \text{PrisonLOS}_T}{(1 + \text{dis})^{\text{PrisonLOS}_T}}$$

The variable dis is the discount rate used in the overall benefit-cost analysis. The data source for Washington State for the variables in (G2.11) are described in Appendix D2.

Change in Post-Jail Community Supervision Costs. The change in post-jail supervision costs, given the above inputs is then computed with:

$$(G2.12) \Delta PostJailCS\$_T = \sum_{y=1}^{PostJailCSLOS_T} \frac{\Delta RC_T \times JailProb_T \times CS\$}{(1 + dis)^{y-1}}, \text{ if } PostJailCSLOS_T < 1, \text{ then } y = PostJailCSLOS_T$$

The variable *dis* is the discount rate used in the overall benefit-cost analysis. The data source for Washington State for the variables in (G2.12) are described in Appendix D2.

Change in Police Costs. The change in police costs is computed by first developing a weighted average per-arrest police cost. Equation (G2.13) computes this estimate for all types of crime *T*. This tool assumes one arrest per prison or jail sentence.

$$(G2.13) Police\$_T = \frac{\sum_{c=1}^{Ct} Police\$_c \times (PrisonSentences_c + JailSentences_c)}{\sum_{c=1}^{Ct} (PrisonSentences_c + JailSentences_c)}$$

$$(G2.14) \Delta Police\$_T = \Delta RC_T \times (PrisonProb_T + JailProb_T) \times Police\$_T$$

The data source for Washington State for the variables in (G2.13) and (G2.14) are described in Appendix D2.

Change in Court Costs. The change in court costs, which includes court personnel, prosecutors, and defenders, given the above inputs is then computed with:

$$(G2.15) Court\$_T = \frac{\sum_{c=1}^{Ct} Court\$_c \times (PrisonSentences_c + JailSentences_c)}{\sum_{c=1}^{Ct} (PrisonSentences_c + JailSentences_c)}$$

$$(G2.16) \Delta Court\$_T = \Delta RC_T \times (PrisonProb_T + JailProb_T) \times Court\$_T$$

The data source for Washington State for the variables in (G2.15) and (G2.16) are described in Appendix D2.

Change in State Government Fiscal Costs

$$(G2.17) \Delta StateFiscal\$ = \Delta Prison\$_T \times PrisonStatePct + \Delta PostPrisonCS\$_T \times PostPrisonStatePct + \Delta Jail\$_T \times JailStatePct + \Delta PostJailCS\$_T \times PostJailCSStatePct + \Delta Police\$_T \times PoliceStatePct + \Delta Court\$_T \times CourtStatePct$$

Change in Local Government Fiscal Costs

$$(G2.18) \Delta LocalFiscal\$ = \Delta Prison\$_T \times (1 - PrisonStatePct) + \Delta PostPrisonCS\$_T \times (1 - PostPrisonStatePct) + \Delta Jail\$_T \times (1 - JailStatePct) + \Delta PostJailCS\$_T \times (1 - PostJailCSStatePct) + \Delta Police\$_T \times (1 - PoliceStatePct) + \Delta Court\$_T \times (1 - CourtStatePct)$$

G3. Estimating the Effect of a Portfolio of Programs That Reduce Crime (Step 3 of 5)

The logic behind WSIPP's sentencing tool is that if a state decides to change its average daily prison incarceration rate, then two things will happen initially: crime levels will change as will net taxpayer costs. Steps 1 and 2 describe the procedures we use to estimate these initial values from a sentencing policy change. The third step in WSIPP's approach then estimates the crime and fiscal effects if a portion, or all, of the net taxpayer savings (from Step 2) is used to fund a portfolio of evidence-based programming resources.

This section describes the inputs needed to create the portfolio of evidence-based crime-reduction programs. It draws on WSIPP's approach and meta-analysis, as described in Appendices A and B. There are three steps to WSIPP's programming portfolio tool.

Review of the Evidence on What Works (and What Does Not). The first step is to produce an estimate of what works and what does not to reduce crime. We begin by analyzing all high-quality research from anywhere in the United States and

elsewhere to determine what options have best achieved desired outcomes (and which have not). Our empirical approach for this first task is to assess systematically, using a meta-analytic framework, the research literature on a given topic. We have found that a meta-analytic approach is particularly helpful in a “real world” policy environment, because it considers all available evidence, not just one or two selected studies. A well done meta-analysis produces an expected effect of a public policy option (and standard error), given the weight of all the evidence. For this analysis, we have examined programs that have achieved reduction in crime outcomes. We report here the results of various types of programs for adult and juvenile offenders. There is also evidence that some prevention programs, such as early childhood education, can reduce crime. We do not show the results of the prevention programs in this document; we are updating reviews of this prevention literature at the present time.

Compute the Economics (Costs and Benefits) of Specific Policy Options. The product of the meta-analyses reveals whether a given programming option can affect crime outcomes. Once this mean effect size (and standard error) is estimated, we bring economics into play by answering two basic questions: “How much does it cost to produce the effect,” and “How much is it worth to people in a state (Washington, in our case) to achieve it?” We have built formal economic models with a consistent set of inputs to measure these costs and benefits. Our “crime model” is discussed in Appendix D2. We present these estimates using standard financial statistics that summarize the cash flows of investments: net present values, benefit-cost ratios, and returns on investment. The analyses provide an internally consistent set of estimates given the estimated effect sizes, the modeling parameters selected, and the modeling structure employed. We present the estimates from three perspectives: the direct participants in policy options; a taxpayer-only perspective; and the non-taxpayer perspective of people who are not the direct program recipient. In the case of crime outcomes, the later perspective is that of crime victims. If crime is reduced, then there is value to people who do not become the victims of crime. If crime is increased, then victimization costs are incurred. The combination of these three perspectives enables a “total state” perspective. For crime outcomes, the taxpayer perspective if further divided into state taxes and local taxes; this information is used in the tool.

Create a Portfolio of Programs Available for Selection. For the sentencing tool, we use our benefit-cost model to create a set of programs that includes adult corrections programs and juvenile justice programs. In the table below, we report the results of our crime model. These results are also displayed on the software screen as shown in Exhibit G1.b. The table displays the number of program evaluations (from the meta analysis) used to draw conclusions about program effectiveness in reducing crime. The table also shows the estimated program cost per participant in 2008 dollars. These costs are on an “intent-to-treat” basis, to be consistent with how we compute crime outcomes. That is, we count both the average costs and the average crime effects of people who start a program, not just those who finish a program. This is done to help avoid the statistical selection bias that can otherwise cloud the causal interpretation of program effectiveness.

We then show the results, per average program participant, of the mean victimizations avoided (and standard error) estimated from the program effect size and other parameters. This process is described in Appendix D2. We then divide the mean victimizations avoided by program costs to show the victimization avoided per dollar of cost. The benefit-cost model described in Appendix D2 estimates the monetary value of the avoided crimes in terms of reduced criminal justice system spending. These results are shown on the table for both mean values and associated standard errors (computed with Monte Carlo simulation after varying a number of parameters). We also compute and report the estimated state portion of the taxpayer benefits. Finally, we indicate the monetary values we compute for the avoided victimization costs.

Exhibit G3.a

Portfolio of High Return-on-Investment Adult Offender and Juvenile Offender Programs to Reduce Crime Results from WSIPP Benefit-Cost Analysis (September 2011), 2008 Dollars

| Program Name | Number of Studies Meta Analyzed | Program Cost Per Program Participant | Victimizations Avoided Per Program Participant | | | Taxpayer Benefits Per Program Participant | | | Victim Benefits Per Program Participant |
|---|---------------------------------|--------------------------------------|--|----------------|--------------------------|---|----------------|---------------------------|---|
| | | | Mean | Standard Error | Per \$/1000 Program Cost | Mean | Standard Error | State Percent of Benefits | |
| Adult Programs | | | | | | | | | |
| Vocational Education in Prison | 3 | \$1,536 | 0.46 | 0.09 | 0.30 | \$4,906 | \$703 | 47% | \$12,569 |
| Education in Prison (basic or post-secondary) | 11 | \$1,102 | 0.49 | 0.10 | 0.45 | \$4,963 | \$1,148 | 47% | \$5,961 |
| Cognitive Behavioral Programs in Prison | 36 | \$217 | 0.26 | 0.11 | 1.19 | \$2,679 | \$1,103 | 47% | \$5,100 |
| Correctional Industries in Prison | 9 | \$1,387 | 0.15 | 0.05 | 0.11 | \$1,546 | \$487 | 47% | \$4,592 |
| Drug Treatment in Prison | 21 | \$3,893 | 0.35 | 0.06 | 0.09 | \$3,459 | \$701 | 47% | \$4,592 |
| Drug Treatment in Community | 6 | \$2,102 | 0.30 | 0.08 | 0.14 | \$3,671 | \$919 | 42% | \$4,804 |
| Drug Courts (adults) | 67 | \$4,095 | 0.20 | 0.04 | 0.05 | \$2,511 | \$275 | 44% | \$4,376 |
| Juvenile Programs | | | | | | | | | |
| Multi-dimensional Treatment Foster Care | 3 | \$7,730 | 0.73 | 0.73 | 0.09 | \$7,747 | \$5,929 | 50% | \$24,068 |
| Family Integrated Transitions | 1 | \$10,993 | 0.54 | 0.24 | 0.05 | \$5,681 | \$2,066 | 50% | \$13,050 |
| Coordination of Services | 1 | \$386 | 0.07 | 0.11 | 0.17 | \$786 | \$1,216 | 45% | \$2,135 |
| Functional Family Therapy | 8 | \$3,191 | 0.74 | 0.34 | 0.23 | \$6,305 | \$2,621 | 45% | \$20,623 |
| Aggression Replacement Training | 4 | \$1,476 | 0.68 | 0.58 | 0.46 | \$6,419 | \$5,053 | 45% | \$9,731 |
| Multi-systemic Therapy | 10 | \$7,206 | 0.63 | 0.42 | 0.09 | \$5,320 | \$3,520 | 45% | \$11,027 |

A user can then select any of these resources to be elements of an overall portfolio. For example, a user could designate 50 percent of the portfolio into cognitive behavioral therapy (CBT) for adult offenders; 25 percent into the juvenile justice program Functional Family Therapy; and 25 percent into the juvenile justice program Aggression Replacement Therapy. The weighted average portfolio effect on crimes avoided (and standard error) as well as the weighted average portfolio cost and taxpayer and victim savings are then computed. The portfolio standard error on the crimes avoided assumes no correlation among the programs selected for the portfolio. The results of a sample portfolio are displayed on the screen shot shown in Exhibit G1.b.

G4. Combinations of Policies (Step 4 of 5)

At this stage, the user has identified a prison ADP change and elasticity assumptions (Step 1); calculated the initial change in criminal justice system costs or benefits from the ADP change's effect of crime (Step 2); and selected a portfolio of evidence-based programs to purchase (Step 3). In Step 4, the user decides on the percentage of taxpayer savings from Step 2 to apply to buying high return on investment program "slots."

Number of Initial Purchased Programming Slots from the Sentencing Change. The number of slots that can be purchased depends initially on the net fiscal cost savings, $\Delta NetFiscal\$$, determined in Step 2. This amount is multiplied by the percentage of funds that will be applied to the initial purchases of evidence-based programming, $InitialPctProgramming$. The user enters this programming percentage in the tool as shown on the screen shot in Exhibit G1.b. The product of the two terms is then divided by the weighted average portfolio cost of the programs purchased, $AvgCostofPortfolioPrograms$, to determine the number of initial slots purchased with the net fiscal savings from the prison ADP reduction. The source of the net fiscal savings is dependent upon whether the user has chosen to decrease or increase prison ADP. For example the net fiscal savings for an ADP reduction results from the direct change in prison expenditures. Alternatively, the net fiscal savings for an increase in ADP derives from the changed costs of criminal justice system if the user has negative elasticities in the tool.

$$(G4.1) \text{InitialSlotsPurchased} = \frac{(\Delta NetFiscal\$ \times InitialPctProgramming)}{AvgCostofPortfolioPrograms}$$

Taxpayer Benefits From the Initial Slots Purchased. The evidence-based slots purchased can be expected to generate taxpayer (and victim) benefits. The weighted portfolio per-participant estimate of these benefits, $AvgTaxpayerBenefitofPortfolio$, is multiplied by the number of initial slots purchased to estimate the total taxpayer benefits expected for the initial slots purchased, $InitialTaxpayerBenefits$. The taxpayer benefits of the evidence-based programs are computed exogenously with WSIPP's benefit-cost model described in Appendix D2.

$$(G4.2) \text{InitialTaxpayerBenefits} = \text{InitialSlotsPurchased} \times \text{AvgTaxpayerBenefitofPortfolio}$$

Cost of Additional Slots Purchased. The expected taxpayer benefits from the initial purchase of evidence-based programs can allow a state to purchase additional evidence-based program slots. The user also enters the parameter for *AdditionalPctProgramming*. The total number of any additional slots is given by:

$$(G4.3) \text{ AdditionalSlotsPurchased} = \frac{(\text{InitialTaxpayerBenefits} \times \text{AdditionalPctProgramming})}{\text{AvgCostofPortfolioPrograms}}$$

Total Slots Purchased. Equation (G4.4) then sums the initial and additional evidence-based portfolio slots purchased to arrive at the total number of slots purchased, *TotalSlotsPurchased*.

$$(G4.4) \text{ TotalSlotsPurchased} = \text{InitialSlotsPurchased} + \text{AdditionalSlotsPurchased}$$

Number of Crimes Avoided With Purchased Programs. The expected number of crimes avoided, *AvoidedCrimes*, is the product of the number of slots and the average crimes avoided per slot, *AvgCrimesAvoidedPerSlot*.

$$(G4.5) \text{ AvoidedCrimes} = \text{TotalSlotsPurchased} \times \text{AvgCrimesAvoidedPerSlot}$$

Net Change in Crime. The net change in crime, *CrimeChange*, is determined by subtracting the avoided crimes as a result of the combination of evidence-based policies from the increase in crimes from the prison ADP reduction, ΔC_t , (from equation G1.3).

$$(G4.6) \text{ CrimeChange} = \Delta C_t - \text{AvoidedCrimes}$$

As described thus far, total crimes avoided are estimated as felony crimes for an average offender. The total crime estimates, however, can be analyzed as violent, property, and drug crimes. We will investigate whether it is possible to enhance the tool to contain this capability.

G5. Risk Analysis (Step 5 of 5)

Analyzing these policy tradeoffs involves a substantial amount of uncertainty. While there is an increasingly strong evidentiary base of knowledge about what works to reduce crime, there remains a considerable level of variation in particular estimates. To reflect this uncertainty, the fifth step in our modeling approach is designed to estimate the riskiness of any combination of policy options.

As with any investment decision, an investor typically wants to know the expected gain of an investment along with a measure of the risk that the investment strategy could produce an undesired result. WSIPP's tool is structured to provide this type of investment information. The bottom-line investment statistics that the WSIPP tool produces include the expected change in taxpayer spending for a portfolio of policy options, along with the risk that the mix of options could lead to more crime, not less.

We estimate the known uncertainty surrounding many of the inputs to the tool. We implement a Monte Carlo simulation approach in Microsoft Excel to vary the following key inputs in the tool. Each time the tool is run (the default is 10,000 cases per run), the tool draws randomly from the user-specified probability distributions for the variables shown in the following table.

**Exhibit G5.a
Parameters Allowed to Vary in Monte Carlo Simulation of the Tool**

| Model Parameter Allowed to Vary | Type of Probability Distribution * |
|---|---------------------------------------|
| Crime-Prison Elasticity and Adjustments | |
| UCR Crime-Prison Elasticity | Triangular |
| Drug Conviction-Prison Elasticity | Triangular |
| UCR Crime-Prison Simultaneity Adjustment | Triangular |
| Drug Conviction-Simultaneity Adjustment | Triangular |
| UCR Crime-Policy Adjustment | Triangular |
| Drug Conviction- Policy Adjustment | Triangular |
| Portfolio Program Effectiveness Estimates | |
| Victimization Avoided Per Program Portfolio Slot | Normal |
| Taxpayer Savings Per Program Portfolio Slot | Normal |
| Additional Indirect variation on Program Portfolio Slots | |
| Program Effect Size | Normal |
| Victimization Costs | Triangular |
| Operating Criminal Justice System Costs | Triangular |
| Discount Rate | Triangular |

* The specific parameters for these distributions are selected by the user. We show WSIPP's values on the screen in Exhibit G1.a.