Benefit-Cost Technical Documentation

Washington State Institute for Public Policy Benefit-Cost Model

June 2016

This Technical Documentation describes the computational procedures used in the Washington State Institute for Public Policy's Benefit-Cost Model.
The Technical Document is periodically updated to incorporate the latest revisions to the model.

Many WSIPP employees, both current and former, contributed to the information contained in the Benefit-Cost Technical Documentation. For further information on the procedures described in this report, contact Stephanie Lee at stephanie.lee@wsipp.wa.gov or (360) 586-3951.

Suggested citation: Washington State Institute for Public Policy (June 2016). *Benefit-cost technical documentation.* Olympia, WA: Author.

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

The Washington State Legislature created the Washington State Institute for Public Policy (WSIPP) in 1983. WSIPP is governed by a Board of Directors that represents the legislature, the governor, and public universities. WSIPP's mission is to carry out practical, non-partisan research at the direction of the legislature or the Board of Directors.

WSIPP's Board of Directors authorized WSIPP to receive outside funding from the MacArthur Foundation and the Pew Charitable Trusts to support some of this work.

Table of Contents

Chapter 1: Overview of the Benefit-Cost Approach and Model 6

1.1 Structure of the Model 7

1.2 General Characteristics of WSIPP's Approach to Benefit-Cost Modeling 8

1.3 Peer Review of the WSIPP Benefit-Cost Model 9

Chapter 2: Procedures to Estimate Effect Sizes and Standard Errors 11

2.1 Effect Sizes from Two Bodies of Research: Program Evaluations and Studies Measuring Linkages Between Outcomes 12

2.2 Meta-Analytic Procedures: Study Selection and Coding Criteria 13

2.3 Meta-Analytic Procedures: Calculating "Unadjusted" Effect Sizes 14

2.4 WSIPP Adjustments to Effect Sizes from Program Evaluations 20
- 2.4a Methodological Quality 21
- 2.4b Researcher Involvement in the Program's Design and Implementation. 21
- 2.4c Evaluations with Weak Outcome Measures. 22
- 2.4d Evaluations Conducted in "Non-Real-World" Settings. 22
- 2.4e Evaluations with Wait-List Research Designs. 22
- 2.4f Values of the Five WSIPP Adjustment Factors. 22
- 2.5g Calculating the Inverse Variance Weights and Standard Errors when WSIPP Adjustments are made to Effect Sizes 24

2.5 WSIPP Adjustments to Effect Sizes from Longitudinal Linkage Studies 24
- 2.5a Methodological Quality 24
- 2.5b Generalizability of the Sample 25
- 2.5c Relevance of the Independent and Dependent Variables 25

2.6 Meta-Analytic Procedures: Calculating "Adjusted" Effect Sizes for the Benefit-Cost Model 25

2.7 The Persistence of Effect Sizes over Time 25

Chapter 3: Procedures to Compute "Monetizable" Outcome Units from Effect Sizes 28

3.1 Effect Size Parameters from Program Evaluations 28

3.2 Monetizable Unit Changes from Effect Sizes from Program Evaluations 30

3.3 Linked Effect Size Parameters 32

3.4 Unit Changes from Linked Effect Sizes 32

3.5 Monetizable Unit Changes for Benefit-Cost Calculation, When a Linked Outcome is Present 33

Chapter 4: Procedures to Estimate Monetary Benefits of Outcome Units 34

4.1 Valuation of Labor Market Outcomes 34
- 4.1a Calculating Earnings 35
- 4.1b Earnings by Educational Attainment 38
- 4.1c Earnings by Population Characteristics 39
- 4.1d Earnings and Employment used in Modeling Disease and Disorder 40
- 4.1e Household Production 40

4.2 Valuation of Crime Outcomes 41
- 4.2a Per-Unit Crime Costs 42
- 4.2b Criminological Information for Different Populations 66

	4.2c	Estimates of Victimizations per Conviction	69
	4.2d	Procedures to Estimate Criminal Justice System and Victimization Events	70
	4.2e	Linkages: Crime and Other Outcomes	72
	4.2f	Special Calculations for Prison and Policing Resources	72
4.3		Valuation of Child Abuse and Neglect Outcomes	80
	4.3a	CAN Prevalence and Cost Parameters	80
	4.3b	Deaths Attributed to CAN	86
	4.3c	Linkages: CAN and Other Outcomes	87
4.4		Valuation of Alcohol, Illicit Drug, and Regular Tobacco Use Outcomes	87
	4.4a	ATOD Epidemiological Parameters: Current Prevalence for Prevention and Intervention Programs	89
	4.4b	ATOD Attributable Deaths	93
	4.4c	Linkages: ATOD and Other Outcomes	96
	4.4d	Human Capital Outcomes Affecting Labor Market Earnings via ATOD-Caused Morbidity and Mortality	96
	4.4e	Medical Costs, Treatment Costs, and Other Costs From ATOD	98
	4.4f	Early Initiation of ATOD	105
4.5		Valuation of Teen Birth Outcomes	105
4.6		Valuation of Public Assistance Outcomes	106
	4.6a	Cash Assistance	106
	4.6b	Food Assistance	106
4.7		Valuation of K–12 Education Outcomes	107
	4.7a	Education Parameters	108
	4.7b	Linkages: Education	110
	4.7c	Valuation of Earnings from High School Graduation	113
	4.7d	Valuation of Earnings from Increases in K–12 Standarized Student Test Scores	116
	4.7e	Valuation of Earnings from Increases in the Number of Years of Education Achieved	117
	4.7f	Valuation of Changes in the Use of K–12 Special Education and Grade Retention	117
	4.7g	Adjustment Factors for Decaying Test Score Effect Sizes to Age 17	118
4.8		Valuation of Mental Health Outcomes	122
	4.8a	Mental Health Parameters.	123
	4.8b	Mental Health Epidemiological Parameters	123
	4.8c	Mental Health Attributable Deaths	124
	4.8d	Linkages: Mental Health to Other Outcomes	126
	4.8e	Human Capital Outcomes Affecting Labor Market Earnings via Mental Health Morbidity and Mortality	126
	4.8f	Medical Costs	128
4.9.		Valuation of Health Care Outcomes	136
	4.9a	Health Care Parameters	137
	4.9b	Health Epidemiological Parameters	140
	4.9c	Deaths Attributable to Health Conditions	143
	4.9d	Linkages: Health Outcomes to Other Outcomes	143
	4.9e	Human Capital Outcomes Affecting Labor Market Earnings via Health Morbidity and Mortality	143
	4.9f	Medical Costs for Specific Health Conditions	146
4.10		Valuation of Labor Market Outcomes	149
4.11		Other Parameters	150
	4.11a	Base Year for Monetary Denomination	150
	4.11b	Discount Rates	150
	4.11c	Demographic Information	151

4.11d Valuation of Reductions in Mortality Risk: Value of a Statistical Life 154
4.11e Deadweight Cost of Taxation 156
4.11f Inflation/Price Indexes 157
4.11g Tax Rates 159
4.11h Capital Costs 159

4.12 Value of an Outcome 159

Chapter 5: Procedures to Avoid Double Counting Benefits 162

5.1 Trumping Rules 1, 2, and 3: When Multiple Outcome Measures Reflect the Same Underlying Construct 162

5.2 Trumping Rule 4: When There are Direct and Indirect Paths to the Same Outcome 163

Chapter 6: Procedures to Estimate Program Costs 164

Chapter 7: Procedures to Access Risk and Uncertainty 166

7.1 Key Inputs Varied in the Monte Carlo Simulation Analysis 166

7.2 Computational Procedures to Carry Out the Simulation 167

Chapter 8: The WSIPP Portfolio Tool 168

8.1 Estimating the Expected Benefits and Costs of Programs and Policies 168

8.2 Preparing Programs and Policies for Portfolio Analysis 168

8.3 Combining Results of a Portfolio of Programs and Policies 169

8.4 Portfolio Risk Analysis 169

Appendix: Estimates of Linked Relationships between Outcomes 171

Chapter 1: Overview of the Benefit-Cost Approach and Model

This Benefit-Cost Technical Documentation describes the latest version of the Washington State Institute for Public Policy (WSIPP) benefit-cost model. The model is designed to produce, for the Washington Legislature, internally consistent estimates of the benefits and costs of various public policies. WSIPP built its first benefit-cost model in 1997 to determine whether juvenile justice programs that have been shown to reduce crime can also pass an economic test. In subsequent years, as WSIPP received new research assignments from the Washington Legislature, the benefit-cost model was revised and expanded to cover additional public policy topics. As of this writing, the legislature or the WSIPP Board of Directors has asked WSIPP to use the benefit-cost model to identify effective public in the following public policy areas:

- Criminal and juvenile justice
- K–12 and early education
- Child welfare
- Substance abuse
- Mental health
- Public health
- Public assistance
- Employment and workforce development
- Health care
- General prevention

The model described in this Technical Documentation reflects our current approach to computing benefits and costs for this wide array of topics. We update and revise our estimates and methods from time to time. In particular, as we use this model in the policy and budgetary process in Washington State, we frequently adapt our approach to better fit the needs of policymakers. This document reflects the current state of the model (as of the publication date on the title page).

This report does not contain our current benefit-cost estimates for these topics; rather, it describes the procedures we use to compute the results. A complete "clickable" list of our current benefit-cost estimates can be found on the WSIPP website.[1]

The overall objective of WSIPP's model is to produce a "What Works?" list of evidence-based public policy options available to the Washington State Legislature, ranked by return on investment. The ranked list can help policymakers choose a portfolio of public policies that are evidence-based and have a high likelihood of producing more benefits than costs. For example, policymakers in the state of Washington can use WSIPP's results to identify a portfolio of evidence-based policies (prevention, juvenile justice, adult corrections, and sentencing policies) that together can improve the chance that crime is reduced in Washington and taxpayer money is used efficiently.

For each evidence-based option we analyze, our goal is to deliver to the legislature two straightforward benefit-cost measures: an expected return on investment and, given the risk and uncertainty that we anticipate in our estimates, the chance that the investment will at least break even (that is, it will have benefits at least as great as costs). To do this, we carry out four basic analytical steps.

1) **What works? What doesn't?** We begin by conducting systematic reviews of the research literature to identify policies and programs that demonstrate an ability to improve specific outcomes. The goal is to assemble all of the best research from around the US (and beyond) that can help inform policymaking in Washington. In Chapters 2 and 3, we describe the methods we use to identify, screen, and code research studies, as well as the meta-analytic approach we use to estimate the expected effectiveness of policy options and to compute "monetizable" units of change.
2) **What passes an economic test?** The second step involves applying economic calculations to put a monetary value on any improved outcome (from the first step). Once monetized, the estimated benefits are then compared

[1] http://www.wsipp.wa.gov/BenefitCost

to the costs of programs or policies to produce an economic bottom line for the investment. Chapter 4 describes the processes we use to monetize the outcomes. Chapter 6 describes our procedures for estimating program costs.

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 involves many individual estimates and assumptions. Almost every modeling step involves at least some level of risk and uncertainty. Chapter 7 describes the "Monte Carlo" approach we use to model this risk. The objective of the risk analysis is to assess the chance that a return on investment estimate will at least break even. 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, the risk question is: given the riskiness in this estimate, what is the chance that the program will at least break even by generating one dollar of benefits for each dollar of cost?

4) **How can a portfolio of evidence-based programs and policies change statewide outcomes?** Finally, the benefit-cost model also allows the user to combine individual policy options into a portfolio. Much like the concept of an investment portfolio in the private sector, this tool allows the user to pick and choose different policy options and project the combined impact of those options on statewide costs, benefits, and outcomes. The WSIPP portfolio tool is the newest aspect of the overall model and is described in Chapter 8.

1.1 Structure of the Model

WSIPP's benefit-cost model is an integrated set of computational routines designed to produce three related benefit-cost summary statistics for each policy option we analyze: a net present value, a benefit-to-cost ratio, and a measure of risk associated with these bottom-line estimates. Each of the summary measures derives from the same set of estimated cash or resource flows over time.

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 1.1.

$$(1.1)\quad \mathrm{NPV_{tage}} = \sum_{\mathrm{y=tage}}^{\mathrm{N}} \frac{\mathrm{Q_y \times P_y - C_y}}{\mathrm{(1 + Dis)^y}}$$

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, multiplied by 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 1.1, Qy, is the estimated number of outcome "units" in year y produced by the program or policy. The procedures we use to develop estimates of Qy are described in Chapters 2 and 3. In Chapter 4 we describe the various methods we use to estimate the price term, Py, in equation 1.1. In Chapter 6 we describe our procedures for computing program costs, Cy. In Chapter 7, 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 equation 1.1, a benefit-to-cost ratio, B/C, can be computed with:

$$(1.2)\quad \frac{B}{C} = \sum_{y=tage}^{N} \frac{Q_y \times P_y}{(1+Dis)^y} \Big/ \sum_{y=tage}^{N} \frac{C_y}{(1+Dis)^y}$$

1.2 General Characteristics of WSIPP's Approach to Benefit-Cost Modeling

Several features are central to WSIPP's benefit-cost modeling approach.

Internally Consistent Estimates. Because WSIPP'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 considered, 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 individual benefit-cost analyses are done, where one researcher conducts an economic analysis for one program and then another researcher performs an entirely different benefit-cost analysis for another program. By adopting one internally consistent modeling approach, our goal is to enable apples-to-apples, rather than apples-to-oranges, benefit-cost comparisons.

Meta-Analysis. The first step in our benefit-cost modeling strategy produces estimates of policies and programs that have been shown to improve particular outcomes. That is, before we undertake any economic analysis of benefits and costs, we first want to determine "what works?" to improve outcomes. To do this, we carefully analyze all high-quality studies from the US (and beyond) to identify well-researched programs or policies that achieve desired outcomes (as well as those that do 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. By including all of the studies in a meta-analysis, we are, in effect, making a statement about the average effectiveness of a particular topic given the weight of the most credible research studies. For example, in deciding whether the juvenile justice 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 of the credible studies we are able to find on Functional Family Therapy.

"Linked" Outcomes. In addition to examining the impacts of a program on directly measured outcomes, we estimate the benefits of linked or indirectly measured outcomes. For example, a program evaluation may measure the direct short-term effect of a child welfare program on child abuse outcomes, but not the longer-term outcomes such as high school graduation. Other substantial bodies of research, however, have measured cause-and-effect relationships between being abused as a child and its effect on the odds of high school graduation. Using the same meta-analytic approach we describe in Chapter 2, we take advantage of this research and empirically estimate the causal "links" between two outcomes. We then use these findings to project the degree to which a program is likely to have longer-term effects beyond those measured directly in program evaluations. The monetization of linked outcomes becomes especially important in conducting benefit-cost analysis when, typically, not all of the impacts of a program are directly measured in the program evaluation studies themselves. We describe how we determine these linkages in Chapter 2, and we list our current estimates for the linkages in this report's Appendix.

Avoiding Double Counting Benefits. We have found that many evaluations of programs and policies measure multiple outcomes. It is desirable, of course, to calculate benefits across multiple outcomes to draw a comprehensive conclusion about the total benefits of a program or policy. To do this, however, runs the risk of double counting outcome measures that really are gauges of the same underlying effect. For example, high school graduation and standardized test scores are two outcomes that may both be measured by a typical program evaluation. These two outcomes, however, are likely to be, at least in part, measures of the same development in a person's human capital. As we describe, we have methods to monetize both outcomes individually and both lead to increased earnings in the labor market. To avoid double counting the benefits of these types of outcomes, we have developed "trumping" procedures, described in Chapter 5.

Measuring Risk. Any tabulation of benefits and costs necessarily involves risk 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 and variances considered. To assess risk, we perform a Monte Carlo simulation technique where we vary the key factors in our calculations. The purpose of the risk analysis is to determine the chance that a particular approach will at least break-even. This type of risk analysis is used by many businesses in investment decision making and we employ the same tools to test the riskiness of public sector options. We describe the Monte Carlo approach in Chapter 7.

Four Perspectives on Benefits and Costs. We categorize estimates of benefits and costs into four distinct perspectives: the benefits and costs that accrue solely to program participants, those received by taxpayers, those received by others, and those that are more indirect.

We created the categories of "Others" and "Indirect" to report results that do not fit neatly in the "participant" or "taxpayer" perspectives. In the "Others" category we include the benefits of reductions in crime victimization, the economic spillover benefits of improvement in human capital outcomes, and payments by private (including employer based) insurers. In the "Indirect" category we include estimates of the net changes in the value of a statistical life and net changes in the deadweight costs of taxation.

The sum of these four perspectives provides a "total Washington" view on whether a program produces benefits that exceed costs. For certain fiscal analyses and state budget preparation, the results of the model can be restricted to focus solely on the taxpayer perspective.

For example, for a juvenile justice program that reduces crime and improves the probability of high school graduation, we record the improved labor market benefits from the increased probability of high school graduation as a participant benefit and the reduced criminal justice system costs from the crime reduction as a taxpayer benefit. In the "Others" category, we include the benefits to crime victims of the reduced crime, along with the economic spillover effects of the high school graduation that accrue to others in society. In the "Indirect" category, we account for the net deadweight costs of taxation (from the costs of the program, as well as the deadweight savings from reduced taxes for future crime avoided).

The Model's Expandability. Evidence-based knowledge is continually expanding. More is known today than ten years ago on the relative effectiveness of programs and still 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 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 flexible design feature allows us to update estimates of the economic bottom lines for public programs. In addition, the model is designed in a modular fashion so that new topic areas (other than those listed in the introduction) can be added to the analysis and modeled in a manner consistent with the topics already analyzed.

1.3 Peer Review of the WSIPP Benefit-Cost Model

WSIPP has had external reviewers examine our work and provide feedback on our methods. In addition, we have had invitations in recent years to publish our work in several peer-reviewed journals.[2]

With assistance from the Pew Charitable Trusts (Pew) and the MacArthur Foundation, WSIPP's benefit-cost model is being implemented in 25 other states as part of the Pew-MacArthur Results First Initiative.[3] As part of our work with these organizations, the benefit-cost model has been reviewed three times in the past six years by an independent team assembled by Pew. Most recently, the benefit-cost model was reviewed in 2014 by:

- D. Max Crowley: NIH Research Fellow, Center for Child & Family Policy, Duke University
- Lynn Karoly: Senior Economist and Director, Office of Research Quality Assurance, RAND Corporation
- David Weimer: Professor of Public Affairs and Political Science, Robert M. La Follette School of Public Affairs, University of Wisconsin-Madison
- Paula Worthington: Senior Lecturer, Harris School of Public Policy, University of Chicago

[2] See: (a) Drake, E. (2012). Reducing crime and criminal justice costs: Washington State's evolving research approach. *Justice Research and Policy, 14*(1), 97-116; (b) Drake, E., Aos, S., & Miller. M. (2009). Evidence-based public policy options to reduce crime and criminal justice costs: Implications in Washington State. *Victims & Offenders: An International Journal of Evidence-based Research, Policy, and Practice*, *4*(2), 170-196; and (c) Lee, S., Drake, E., Pennucci, A., Bjornstad, G., & Edovald. T. (2012). Economic evaluation of early childhood education in a policy context. *Journal of Children's Services*, *7*(1), 53-63.

[3] See: http://www.pewstates.org/projects/pew-macarthur-results-first-initiative-328069

The benefit-cost model was also reviewed in 2012 by Kirk Jonas (Director, Office of Research Compliance and Integrity, University of Richmond, Virginia), Steven Raphael (Professor of Public Policy, Goldman School of Public Policy, University of California-Berkeley), Lynn Karoly, and David Weimer, and in 2010 by David Weimer, Lynn Karoly, and additionally, Mike Wilson (Economist, Oregon Criminal Justice Commission).

Annually between 2011 and 2015, Pew hosted meetings with the states involved in the Pew-MacArthur Results First Initiative. Approximately 50-100 participants attended each of the annual meetings. During this time, WSIPP received questions, comments, and criticisms on our technical and non-technical aspects of our methods, software, and policy scenarios. These observations have been helpful to us as we update the model.

Lastly, Pew has technical assistance consultants responsible for learning the benefit-cost model in order to assist the states in implementing the model. The technical assistance consultants have been using the benefit-cost model since 2010, and continually provide feedback on our approach.

Building a far-reaching benefit-cost model requires many modeling decisions. Our choices are not necessarily the ones that all of the reviewers would have made. Thus, while we have benefited from all of the comments, we remain solely responsible for our modeling choices.

Chapter 2: Procedures to Estimate Effect Sizes and Standard Errors

As outlined in Chapter 1, the WSIPP model is an integrated set of computational routines designed to produce internally consistent benefit-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 2.1.

$$(2.1)\quad NPV_{tage} = \sum_{y=tage}^{N} \frac{Q_y \times P_y - C_y}{(1+Dis)^y}$$

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, multiplied by 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 2.1, Q_y, is the estimated number of outcome "units" in year y produced by the program or policy. The Q_y, term in equation 2.1 is, in turn, a function of two factors in the WSIPP model: an "effect size" (ES) and a "$Base$" variable as given by equation 2.2.

$$(2.2)\quad Q_y = f(ES, Base)$$

The WSIPP model is designed to accommodate outcomes that are measured either with continuous scales (e.g. standardized student test scores) or as dichotomies (e.g. high school graduation).

For continuously measured outcomes, as given by equation 2.3 and described later in this Chapter and in Chapter 3, Q_y is calculated with a Cohen's d effect size and a $Base$ variable, which is measured as a standard deviation of the outcome measurement.

$$(2.3)\quad Q_y = Base_y \times ES$$

For dichotomously measured outcomes, Q_y is calculated with a D-cox effect size and a $Base$ variable, which is measured as a percentage. Our precise procedures to calculate Q_y for dichotomies are discussed in Chapter 3, but the essential procedure follows equation 2.4.[4]

$$(2.4)\quad Q_y = \frac{\left(e^{ES\times 1.65} \times Base_y\right)}{\left(1 - Base_y + Base_y \times e^{ES\times 1.65}\right)} - Base_y$$

Two exceptions to this equation for estimating Q_y for continuously measured outcomes are when: (1) an effect size is measured via percent change or "semi-elasticity" in an outcome (currently, WSIPP uses this method for direct labor market earnings measured by workforce development programs and health care costs and frequency of visits measured by evaluations of certain health care programs), (2) an effect size is measured via an elasticity, currently used for certain measures of crime and certain measures of health care costs. For these conditions, we use equation 2.5 below.

$$(2.5)\quad Q_y = ES$$

This Chapter describes the process we use to estimate the effect size term, ES, in equations 2.3 and 2.4. Chapter 3 discusses how Q_y is then estimated from the effect sizes and dichotomous or continuous base variables. In Chapter 4 we

[4] The D-cox transformation that we employ, as well as other possible transformations of dichotomous data to approximate a standardized mean difference effect size, produces results that are known to introduce distortions when base percentages are either very large or very small. The D-cox has been shown to introduce fewer distortions than other procedures, but the D-cox remains problematic when base rates are very low or high. See: 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. In Chapter 3 we describe our current procedures designed to reduce these distortions.

describe the various methods we use to estimate the price term, P_y, in equation 2.1. In Chapter 6 we describe our procedures for computing program costs, C_y, in equation 2.1.

2.1 Effect Sizes from Two Bodies of Research: Program Evaluations and Studies Measuring Linkages Between Outcomes

To estimate the effect of a program or policy on outcomes of interest, WSIPP's approach draws on two bodies of research. First, we compute effect sizes from program evaluation research; this type of research measures whether a program or policy has a causal effect on outcomes of interest. Second, to supplement and extend the program evaluation research, we use other bodies of evidence that examine causal "linkages" between two different outcomes. The overall goal is to combine the best current information from these two bodies of research to derive long-run benefit-cost estimates for program and policy choices.

The logic of using "linkage" studies to support program evaluation findings follows the path illustrated in this expression:

$$if\ Program \rightarrow Outcome_1, \quad and\ if\ Outcome_1 \rightarrow Outcome_2, \quad then\ Program \rightarrow Outcome_2$$

That is, if a meta-analysis of program evaluations—the first body of research—establishes a causal effect of a program ($Program$) on one outcome ($Outcome_1$), and another body of linkage research measures a causal temporal relationship between that outcome ($Outcome_1$) and another outcome ($Outcome_2$) of interest, then it logically follows that the program is likely to have an effect on the second outcome, in addition to having an effect on the directly measured first outcome.

These relationships are important for benefit-cost analysis because, unfortunately, many program evaluations do not measure all of the longer-term outcomes of interest. For example, we have meta-analyzed all credible program evaluations of a juvenile justice program called Functional Family Therapy (FFT) and found that the program reduces juvenile crime—the first step in the expression above. Crime is an important outcome and it is measured in the program evaluations of FFT. We label this a "directly" measured outcome since it was estimated in the program evaluations themselves.

Unfortunately, the outcome evaluations of FFT did not measure whether the program affects high school graduation rates—a second outcome of keen interest to the Washington Legislature. There are, however, other substantial bodies of longitudinal research that indicate how changes in one outcome causally lead to changes in a second outcome. For example, we have separately meta-analyzed credible longitudinal research studies that identify a causal relationship between juvenile crime and high school graduation—the second step in the expression above. We label this relationship a "linked" outcome since it was not estimated in the FFT evaluations themselves, but can be reasonably inferred by applying the results of other credible longitudinal research. We list our current estimates for the linkages in this report's Appendix.

Therefore, we compute effect sizes and standard errors, with the procedures described below, for both direct and linked outcomes and we use them in our benefit-cost analysis.

2.2 Meta-Analytic Procedures: Study Selection and Coding Criteria

To estimate the effects of programs and policies on outcomes, we employ statistical procedures researchers have developed to facilitate systematic reviews of evaluation evidence. This set of procedures is called "meta-analysis."[5] A meta-analysis—sometimes referred to as a "study of studies"—produces a weight-of-the-evidence summary of a collection of individual program evaluations (or studies of the longitudinal relationships between outcomes) on a given topic. The general idea is to 1) define a topic of interest (e.g. do drug courts lower crime; does child abuse and neglect reduce the probability of high school graduation?), 2) gather all of the credible evaluations that have been done on the topic from around the US and beyond, and 3) use meta-analysis to draw an overall conclusion about the average effectiveness of a program to achieve a specific outcome or the relationship between one outcome and another.

A meta-analysis is only as good as the selection and coding criteria used to conduct the study.[6] Following are the key choices we implement.

Study Selection. We use four primary means to locate studies for meta-analysis of programs: 1) we consult the bibliographies of systematic and narrative reviews of the research literature in the various topic areas; 2) we examine citations in the individual studies we locate; 3) we conduct independent literature searches of research databases using search engines such as Google, Proquest, Ebsco, ERIC, PubMed, and SAGE; and 4) we contact 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 is that an evaluation must either have a control or comparison group or use advanced statistical methods to control for unobserved variables or reverse causality. If a study appears to meet these criteria, we then secure a copy of the study for our review.

Peer-Reviewed and Other Studies. We examine all evaluation studies we can locate with these search procedures. Many of these studies are published in peer-reviewed academic journals while others are from reports obtained from government agencies or independent evaluation contractors. 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 we can locate that meet our criteria, regardless of published source.

Intent-to-Treat Samples. We do not include a study in our meta-analytic review if the treatment group is 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 evaluation studies of program completers, however, also contain information on program dropouts in addition to a comparison group. In these situations, we include the study if sufficient information is provided to allow us to reconstruct an intent-to-treat group that includes both completers and non-completers, or if the demonstrated rate of program non-completion is very small. In these cases, the study still needs to meet our other inclusion requirements.

Random Assignment and Quasi-Experiments. Random assignment studies are preferred for inclusion in our review, but we also include studies with non-randomly assigned comparison groups. We only include quasi-experimental studies if sufficient information is 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. Since we follow the statistical procedures in Lipsey and Wilson,[7] a study must provide the necessary information to calculate an effect size, as described below. If the necessary information is not provided, and we are unable to obtain the necessary information directly from the study's author(s), the study is not included in our review.

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

[6] All studies used in the meta-analyses for individual programs and policies are identified in the detailed results documented in WSIPP programs, which can be found on the WSIPP website: http://www.wsipp.wa.gov. Many other studies were reviewed, but did not meet the criteria set for this analysis.

[7] Lipsey & Wilson, (2001).

Multivariate Results Preferred. Some studies present two types of analyses: raw outcomes that are not adjusted for covariates such as age, gender, or pre-intervention characteristics; and those that are adjusted with multivariate statistical methods. In these situations, we code the multivariate estimates focusing on the author's preferred specification.

Averaging Effect Sizes for Similar Outcomes so Each Study Contributes One Outcome. Some studies report similar outcomes: e.g., reading and math test scores from different standardized assessments. In such cases, we average the similar measures and use 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. This avoids one study having more weight in a meta-analysis simply because it measured more outcomes.

Outcomes Measured at Different Follow-Up Periods. If outcomes for study samples are measured at multiple points in time, and if a sufficient number of studies contain multiple, similar follow-up periods, we calculate effect sizes for both an initial and longer term follow-up periods. Using different points of time of measurement allows us to examine, via meta-regression, whether program effects change (i.e., decay or increase) over time.

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

- **Two-tail p-values.** Some studies only report p-values for significance testing of program outcomes. When we have to rely on these results, if the study reports a one-tail p-value, we convert it to a two-tail test.
- **Declaration of significance by category.** Some studies report results of statistical significance tests in terms of categories of p-values. Examples include: p<0.01, p<0.05, or non-significant at the p=0.05 level. We calculate effect sizes for these categories by using the highest p-value in the category. Thus, if a study reports significance at p<0.05, we calculate the effect size at p=0.05. This is the most cautious strategy. If the study simply states a result is non-significant, but does not indicate a p-value, then we load in a zero effect size, unless some other piece of information reported in the study (perhaps a graph) provides some indication of the direction of the effect, in which case we compute the effect size assuming a p-value of 0.50.

2.3 Meta-Analytic Procedures: Calculating "Unadjusted" Effect Sizes

Effect sizes summarize the degree to which a program or policy affects an outcome, or the degree that one outcome is causally related to another outcome. In experimental program settings this involves comparing the outcomes of treated participants relative to untreated participants. Analysts use several methods to calculate effect sizes, as described in Lipsey and Wilson.[8] The most common effect size statistic, and the measure we use in our meta-analyses, is the standardized mean difference effect size.

Continuously Measured Outcomes. The mean difference effect size is 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 the following equation:

$$(2.6)\quad ES = \frac{M_t - M_c}{\sqrt{\dfrac{(N_t - 1)SD_t^2 + (N_c - 1)SD_c^2}{N_t + N_c - 2}}}$$

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.

In many research studies, the numerator in equation 2.6, $M_t - M_c$, 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 equation 2.6 is the standard deviation for the entire sample. In these types of regression studies, unless information is present that allows the

[8] Lipsey & Wilson, (2001).
[9] Ibid, Table B10, equation 1, p. 198.

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_tN_c is set to equal $(N/2)^2$.

We compute the variance of the mean difference effect size statistic in equation 2.6 with the following equation:[10]

$$(2.7)\quad ESVar = \frac{N_t + N_c}{N_tN_c} + \frac{ES^2}{2(N_t + N_c)}$$

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 the following equation:[11]

$$(2.8)\quad ES = t\sqrt{\frac{N_t + N_c}{N_tN_c}}$$

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. shows that the Cox transformation produces the most unbiased approximation of the standardized mean effect size.[12] Therefore, to approximate the standardized mean difference effect size for continuously measured outcomes, we calculate the effect size for dichotomously measured outcomes with the following equation:

$$(2.9)\quad ES_{Cox} = \frac{\ln\left[\frac{P_t(1-P_c)}{P_c(1-P_t)}\right]}{1.65}$$

where P_t is the percentage of the treatment group with the outcome and P_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

$$(2.10)\quad ESVar_{Cox} = .367\left[\frac{1}{O_{1t}} + \frac{1}{O_{2t}} + \frac{1}{O_{1c}} + \frac{1}{O_{2c}}\right]$$

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 report the results of statistical analysis such as chi-square (χ^2) statistics. In these cases, we first estimate the absolute value of $ES_{arcsine}$ per Lipsey and Wilson,[13] then based on analysis we conduct, we multiply the result by 1.35 to determine ES_{Cox} as given by the following equation:

$$(2.11)\quad |ES_{cox}| = 1.35 * 2\sqrt{\frac{X^2}{N_t + Nc - X^2}}$$

Similarly, we determine that in these cases, using equation 2.7 to calculate the variance underestimates $ESVar_{Cox}$ and, hence over-estimates the inverse variance weight. We conducted an analysis which shows that $ESVar_{Cox}$ is linearly related to $ESVar$. Our analysis indicates that multiplying $ESVar$ by 1.77 provides a very good approximation of $ESVar_{Cox}$.

[10] Ibid, Table 3.2, p. 72.
[11] Ibid, Table B10, equation 2, p. 198.
[12] Sánchez-Meca et al., (2003).
[13] Lipsey & Wilson, (2001), Table B10, equation 23, p. 200.

Odds Ratios and Confidence Intervals.

Sometimes authors report dichotomous outcomes as odds ratios and confidence intervals. In those instances we calculate the effect size using equation 2.9, i.e. taking log of the odds ratio divided by 1.65.
The variance is calculated using the following equation:

$$(2.12)\ \ Var_{adj} = 0.367 \left\{\frac{\ln(upper\ CI) - \ln(lower\ CI)}{2}\right\} / 1.96$$

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

Effect Sizes Measured as Elasticities or Semi-elasticities.
Some of the research literatures we review are econometric in nature; that is, they use regression techniques econometricians often use to consider unobserved variables bias or simultaneity. The metric used in many of these economic studies to summarize results when analyzing a continuous outcome is an elasticity—how a percentage change in one continuously measured "treatment" affects the percentage change in a continuously measured outcome—or a semi-elasticity also known as a percent change—how a dichotomously measured "treatment" affects a percent change in a continuously measured outcome. For example, the research literatures that measure the impact of increased incarceration rates on crime and the effects of the number of police officers on crime both use elasticities to describe the relationships. For studies that do not estimate elasticities directly, we compute the elasticity from the author's preferred regression coefficient taken at the study's mean values. Similarly, research estimating the effect of participating in high deductible health care plan on health care costs often use semi-elasticities. We would then estimate a semi-elasticity, or percent change, in health care costs due to participation in a high-deductible plan. Thus, the effect size for these analyses is an elasticity or semi-elasticity, rather than the other effect size metrics (Cohen's D or D-cox effect sizes) used when we conduct meta-analyses of programs.

Modifying Effect Sizes for Small Sample Sizes. Since some studies have very small sample sizes, we follow the recommendation of many meta-analysts and account for this. Small sample sizes have been shown to upwardly bias effect sizes, especially when samples are less than 20. Following Hedges, Lipsey and Wilson 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), as given in the following equation: [14]

$$(2.13)\ \ ES'_m = \left[1 - \frac{3}{4N - 9}\right] *\ \ ES_m$$

Modifying Effect Sizes and Variances for Multi-Level Data Structures. Many studies measure the results of programs that are delivered in hierarchical structures. For example, in the education field, 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 of this sort will underestimate the variance in outcomes and, thus, may over-estimate effect sizes. In studies that do not account for clustering, effect sizes and their variance require additional adjustments.[15]

There are two types of studies, each requiring a different set of adjustments.[16]

[14] Lipsey & Wilson, (2001), equation 3.22, p. 49 and 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] Studies that employ hierarchical linear modeling, fixed effects with robust standard errors, or random effects models account for variance and need no further adjustment.
[16] These formulas are taken from: Hedges, L. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics, 32*(4), 341-370.

First, for individual-level studies that ignore the variance due to clustering, we make adjustments to the mean effect size and its variance, using the following equation:

$$(2.14)\quad ES_T = ES_m * \sqrt{1 - \frac{2(n-1)\rho}{N-2}}$$

$$(2.15)\quad V\{ES_T\} = \left(\frac{N_t + N_c}{N_t N_c}\right)[1 + (n-1)\rho] + ES_T{}^2\left(\frac{(N-2)(1-\rho)^2 + n(N-2n)\rho^2 + 2(N-2n)\rho(1-\rho)}{2(N-2)[(N-2) - 2(n-1)\rho]}\right)$$

where ρ is the intraclass correlation coefficient, 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.

For example, in the educational field, clusters can be classes, schools, or districts. To meta-analyze education studies, we use 2006 Washington Assessment of Student Learning (WASL) data to calculate values of ρ for the school-level (ρ = 0.114) and the district level (ρ = 0.052). Class-level data are not available for the WASL, so we use a value of ρ = 0.200 for class-level studies.

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

$$(2.16)\quad ES_T = ES_m * \sqrt{\frac{1 + (n-1)\rho}{n\rho}} * \sqrt{\rho}$$

$$(2.17)\quad v\{ES_T\} = \left\{\left(\frac{N_t - N_c}{N_t N_c}\right) * \left(\frac{1 + (n-1)\rho}{n\rho}\right) + \frac{[1 + (n-1)\rho]^2 * ES_T{}^2}{2n\rho(K-2)}\right\} * \rho$$

We do not adjust effect sizes in studies reporting dichotomous outcomes. This is because the Cox transformation assumes the entire normal distribution at the student level.[17] However, when outcomes are dichotomous, we use the "design effect" to calculate the "effective sample size."[18] The design effect is given by the following equation:

$$(2.18)\quad D = 1 + (n-1)\rho$$

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 given by the following equation:

$$(2.19)\quad N_{t(eff)} = \frac{N_t}{D}$$

In some studies, for example in a mental health setting where the treatment group receives an intervention (therapy) and the comparison group does not, the treatment group may be clustered within therapists while the comparison group is not clustered. To our knowledge, there are no published methods for corrected effect sizes and variance for such studies. Dr. Larry Hedges provided the following approach for these corrections.

We first calculate an intermediate estimate of ES using the following equation:[19]

$$(2.20)\quad ES_{int} = ES * \sqrt{1 - \frac{m_t(n_t - 2)\rho}{N-2}}$$

where m_t is the number of clusters in the treatment group, and n_t is the number of subjects in the treatment group, and N is the total sample size.

[17] Mark Lipsey (personal communication, November 11, 2007).
[18] 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/
[19] Larry Hedges (personal communication, June 11, 2012).

Then an approximately unbiased estimate of ES_T is obtained by multiplying ES_{int} by $J(h)$, where h is the effective degrees of freedom as given by the following equation:[20]

$$(2.21)\quad h = \frac{[(N-2)(\rho-1)+(n_t m_t - n_t)\rho]^2}{(N-2)(1-\rho)^2+(n_t m_t - n_t)n_t\rho^2+2(n_t m_t - n_t)\rho(1-\rho)}$$

and $J(h)$ is given by the following equation:[21]

$$(2.22)\quad J(h) = 1-\frac{3}{4h-1}$$

Thus, the final unbiased estimate of ES_T is:[22]

$$(2.23)\quad ES_T = ES_{int} * J(h)$$

The variance of the effect size when only one group is clustered is given by the following equation:[23]

$$(2.24)\quad ESVar = \frac{1+(n-1)\rho}{n_t m_t}+\frac{1-\rho}{m_c}+\frac{([N-2)(1-\rho)^2+(n_t m_t - n_t)n\rho^2+2(n_t m_t - n)\rho(1-\rho] * ES_t}{2[(N-2)(1-\rho)+(n_t m_t - n_t)\rho]^2}$$

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 most mean effect sizes is computed with the following equation:[24]

$$(2.25)\quad SE_T = \sqrt{\frac{N_t+N_c}{N_t N_c}+\frac{ES^2}{2(N_t+N_c)}}$$

For effect sizes measured as elasticities or semi-elasticities, the SE_T is equivalent to the standard error of the elasticity or semi-elasticity. When a study reports the standard error on the elasticity or semi-elasticity, we use that value as SE_T. The standard error of the elasticity is most commonly reported when the study estimates the elasticity from a log-log model, or the semi-elasticity from log-level model.

If a study does not report the elasticity standard error, but calculates an elasticity or semi-elasticity from a linear model, we calculate the SE_T from the linear model using the following equations:

For an elasticity from a linear model the variance of the elasticity is calculated as by the following equation:

$$(2.26)\quad Var(Elas) = \frac{X^2}{Y^2}Var(\beta_1)+\beta_1^2 * \frac{X^2}{Y^4} * Var(Y)$$

where β_1 is the coefficient on X. Then, SE_T is the square root of the variance.

[20] Ibid.
[21] Ibid.
[22] Ibid.
[23] Ibid.
[24] Lipsey & Wilson, (2001), equation 3.23, p. 49.

For a semi-elasticity from a linear model, we can calculate the variance with the following equation:

$$(2.27)\ Var(SemiElas\ or\ \%\ Change) = \left(\frac{Y_t}{Y_c}\right)^2 * \left(\frac{Var(Y_c)}{Y_c^2} + \frac{Var(Y_t)}{Y_t^2}\right)$$

where Y_t and Y_c are the Y values for the treatment and comparison groups (e.g. health care expenditures).

Finally, when a standard error is not reported and cannot be calculated from the information provided in the study, we estimate the standard error of the elasticity using the reported t-statistic for the regression coefficient from which the elasticity is estimated. For example, if a study uses the coefficient β to calculate an elasticity, and the t-statistic on β is reported as t_β, we calculate the standard error on the elasticity for that study as shown in the following equation:

$$(2.28)\quad SE_T = \frac{ES_\epsilon}{t_\beta}.$$

Second, for standardized mean difference and elasticity effect sizes, the inverse variance weight w is computed for each mean effect size with the following equation:[25]

$$(2.29)\quad w = \frac{1}{SE_T^2}$$

[For dichotomous outcomes, the inverse variance weight w is computed by taking the inverse of the variance presented in equation 2.10 (on page 15).

The weighted mean effect size for a group with i studies is computed with the following equation:[26]

$$(2.30)\quad \overline{ES} = \frac{\sum(w_i ES_{T_i})}{\sum w_i}$$

Confidence intervals around this mean are then computed by first calculating the standard error of the mean with the following equation:[27]

$$(2.31)\quad SE_{\overline{ES}} = \sqrt{\frac{1}{\sum w_i}}$$

Next, the lower, ES_L, and upper limits, ES_U, of the confidence interval are computed with the following equation:[28]

$$(2.32)\quad \overline{ES_L} = \overline{ES} - z_{(1-\alpha)}\left(SE_{\overline{ES}}\right)$$

$$(2.33)\quad \overline{ES_U} = \overline{ES} + z_{(1-\alpha)}\left(SE_{\overline{ES}}\right)$$

In equations 2.32 and 2.33, $z_{(1-\alpha)}$ is the critical value for the z-distribution (1.96 for α = 0.05).

[25] Ibid., equation 3.24, p. 49. Note that for our IVW calculations, we use the error of the elasticity coefficient the standard error for an elasticity.
[26] Lipsey & Wilson, (2001), p. 114.
[27] Ibid.
[28] Ibid.

The test for homogeneity, which provides a measure of the dispersion of the effect sizes around their mean, is given by the following equation:[29]

$$(2.34)\ Q_i = \left(\sum w_i\, ES_i^2\right) - \frac{\left(\sum w_i\, ES_i^2\right)}{\sum w_i}$$

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, we use a random effects model to calculate the weighted average effect size. Random effects models allow us to account for between-study variance in addition to within-study variance.[30]

This is accomplished by first calculating the random effects variance component, v using the following equation:[31]

$$(2.35)\ v = \frac{Q_i - (k-1)}{\sum w_i - \left(\sum wsq_i / \sum w_i\right)}$$

where wsq_i is the square of the weight of $ES_{i.}$(2.20).

This random variance factor is 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.

2.4 WSIPP Adjustments to Effect Sizes from Program Evaluations

In WSIPP reports and on our website, we show the results of our meta-analyses calculated with the standard meta-analytic formulas described in Chapter 2.3, above. We call these effects "unadjusted effect sizes." In our reports and on our website, we also list an "adjusted effect size" for each topic. These adjusted effect sizes, which are modifications of the unadjusted results, may be smaller, larger, or equal to the unadjusted effect sizes we report. It is important to note that we use the adjusted effect sizes, not the unadjusted effect sizes, in our benefit-cost model.

In this section, we describe our rationale and procedures for making adjustments to the effect size results from program evaluations. The overall goal of WSIPP's benefit-cost model is to supply the Washington State Legislature with information about what will work to improve outcomes in Washington. If a program has been rigorously tried and tested somewhere else, we want to be able to make an inference about whether it is likely to work in Washington. As we detail below, we think there is reason to be concerned that the results of individual program evaluations (the ones we enter into our meta-analyses) may not be obtained if the program is tried in Washington. Many evaluations of program effectiveness occur under conditions that may not reflect what we would expect in real-world implementation in Washington.

Therefore, to better estimate the results we would expect to achieve in Washington, we developed five types of adjustments. As we explain below, if we determine it to be necessary, we make adjustments for:

(a) the methodological quality of each study we include in a meta-analyses;
(b) whether the researcher(s) who conducted a study is (are) invested in the program's design and results;
(c) the relevance or quality of the outcome measured used in a study;
(d) whether the research was conducted in a laboratory or other unusual "non-real world" setting; and
(e) situations in which an evaluation of a program was conducted against a wait-list comparison group, as oppose to a treatment-as-usual comparison group.

[29] Ibid., p. 116.
[30] 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.
[31] Ibid., p. 134.

2.4a Methodological Quality.

Not all research is of equal quality, and this variation has the potential to systematically bias the results of a study. Some studies are able to use "gold standard" research designs that are well implemented, and the results can be viewed as accurate representations of whether or not the program had a causal effect on an outcome. Other studies may not be able to use the best research designs; these studies may reduce the confidence that can be placed in making cause-and-effect inferences. In particular, studies with less rigorous research designs cannot completely control for self-selection bias or other unobserved threats to the validity of the reported evaluation results. This does not mean that results from these studies are of no value; rather, it just means that less confidence can be placed in any cause-and-effect conclusions drawn from the results.

We assign program evaluation studies to different "research design" categories based on their methodology. This categorization allows us, via meta-regression, to account for the degree to which, on average, differences in the quality of research designs may affect a program's true effect on outcomes. As we explain below, we then use this meta-regression information to adjust effect size results, if necessary. We list our current adjustments for research design in Section 2.4f in this document.

The following research design categories are used:

- **Category 5** includes well-implemented random assignment studies in which subjects are assigned to a treatment group and a control group who do not receive the treatment/program. Studies categorized as a 5 must indicate how well the random assignment occurred by reporting values for pre-existing characteristics for the treatment and control groups.
- **Category 4** includes experimental random assignment studies with implementation problems or studies that use a lottery or random assignment approach from a wait-list when programs are oversubscribed. Random assignment studies in this category, for example, could have crossovers between the treatment and control groups or differential attrition rates between the groups.
- **Category 3** includes natural experiments or studies that use advanced methods in an attempt to control for unobserved variables or reverse causality. Studies categorized as a 3 include instrumental-variable approaches, regression discontinuity designs, panel data analyses with fixed effects, difference-in-differences, or a Heckman approach to modeling self-selection.[32]
- **Category 2** includes quasi-experimental research designs where the treatment and comparison groups are reasonably well matched on pre-existing differences in key variables. For this category, studies must demonstrate that few, if any, significant differences are observed in relevant pre-existing variables. Alternatively, an evaluation must employ sound multivariate statistical techniques (e.g., logistic regression, hierarchical linear modeling for nested variables, or propensity score matching) to control for pre-existing differences.
- **Category 1** includes quasi-experimental studies that are less well-implemented or do not use many statistical controls to control for differences between the treatment and control groups.

Program evaluation studies that do not fit into these categories are assigned to "Category 0" which means that they are not included in our meta-analysis because we cannot confidently estimate a causal treatment effect of the program. Categorizing programs with this scheme is, at least to a degree, subjective. We rely on the accumulated experience of WSIPP analysts to make consistent coding decisions about these research design distinctions.

2.4b Researcher Involvement in the Program's Design and Implementation.

As noted, the purpose of the WSIPP's work is to identify programs that can make cost-beneficial improvements to Washington's actual public service delivery system. There is some evidence that programs closely controlled by researchers

[32] 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. Schlotter, M., Schwerdt, G., & Woessman, L. (2011). Econometric methods for causal evaluation of education policies and practices: a non-technical guide. *Education Economics, 19(2),* 109-137.

or program developers have consistently better results than those that operate in "real world" administrative structures.[33] In our own evaluation of a real-world implementation of a research-based juvenile justice program in Washington, we found that the actual results were lower than the results obtained when the intervention was conducted by the originators of the program.[34] Therefore, because we are concerned that effects observed in developer-controlled evaluations may often overstate the effects we might expect in real-world application in Washington, we code each study by noting whether the developer was involved in the program or evaluation. We then may make an adjustment to effect sizes to reflect this distinction. When possible, we use the results of our meta-regressions to inform the magnitude of any adjustment; lacking meta-evidence to compute a topic-specific adjustment empirically, we may make an adjustment based on *a priori* assumptions, which are themselves informed by our previous analyses of other policy topics. We list our current adjustments for developer involvement in Section 2.4f.

2.4c Evaluations with Weak Outcome Measures.

Some evaluations use outcome measures that may not be precise gauges of the ultimate outcome of interest to Washington. In these cases, we record a flag that we can use in a meta-regression to determine if an adjustment is necessary. We list our current adjustments for weak outcome measures in Section 2.4f.

2.4d Evaluations Conducted in "Non-Real-World" Settings.

As noted, the purpose of the WSIPP's assignments from the Washington Legislature is to identify programs that can make cost-beneficial improvements to Washington's actual public service delivery systems. We code each study by noting whether the program was delivered in a "real-world" setting similar to what would occur in Washington, or whether it was done in an unusual setting, such as a university-based experiment. We then may make an adjustment to effect sizes to reflect this distinction. When possible, we use the results of our meta-regressions to inform the magnitude of this adjustment; lacking evidence to compute a topic-specific adjustment empirically, we may make an adjustment based on *a priori* assumptions which are themselves informed by our analyses of other policy topics. We list our current adjustments for non-real-world settings in Section 2.4f.

2.4e Evaluations with Wait-List Research Designs.

In some topic areas, for example, mental health interventions, our goal is to estimate the average effect of a program compared to non-specific treatment as usual. While some program evaluations measure treatment as usual for the comparison group, other studies compare a treatment group to a wait-list or no-treatment group. We find that average effect sizes are smaller when the comparison group is treatment as usual or an attention placebo, compared to no-treatment control groups. Therefore, when our goal is to estimate the effect of a specific treatment vs. treatment as usual, we may make an adjustment to the effect size to reflect the distinction between active comparisons and no treatment, based on meta-regression of studies in similar topic areas.

2.4f Values of the Five WSIPP Adjustment Factors.

As noted, we base the magnitude of our adjustments for each of these five factors on evidence, wherever possible. That is, when there are sufficient number of studies for us to analyze, we conduct meta-regressions (multivariate linear regression analysis, weighted by inverse variances) in a research area to estimate how much of an adjustment (if any) to make for each of these five factors. Lacking enough studies to conduct a topic-specific meta-regression, we may also make adjustments based on our accumulated knowledge about how these factors can be expected to influence whether specific program evaluation results are likely to be applicable to Washington. In such cases, these *a priori* adjustments represent our informed judgments, until they can be replaced with the results of topic-specific meta-regressions.

[33] 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. Lipsey found that, for juvenile delinquency evaluations, programs in routine practice (i.e., "real world" programs) produced effect sizes only 61% 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.
[34] Barnoski, R. (2004). *Outcome evaluation of Washington State's research-based programs for juvenile offenders*. (Doc. No. 04-01-1201). Olympia: Washington State Institute for Public Policy.

To estimate these adjustments, we undertake a series of meta-regression analyses, one for each broad research area. In some cases, where a research literature is particularly large, we may perform these meta-regressions on smaller groups of topics. In each meta-regression, we include all effect sizes included in our meta-analyses for that topic area, weight by the random effects inverse variance for each, and cluster standard errors by each study in the analysis. Our independent variables include the five factors described above.

An explicit adjustment factor (in the form of a multiplier) is assigned to the results of individual effect sizes based on our findings. Adjustments are made by multiplying the unadjusted effect size for any study by the adjustment factors for the topic area. The resulting meta-analytic results for the adjusted effect sizes are then used in the benefit-cost analysis, as explained in Section 2.6.

The following table lists the current WSIPP adjustments—in the form of multiplicative factors applied to unadjusted effect sizes—that we apply to broad topic areas for the five factors.

Exhibit 2.0

Current WSIPP Adjustments—in the Form of Multiplicative Factors Applied to Unadjusted Effect Sizes

Topic area	Research design	Researcher = developer	Weak outcome measure	"Not real world"	Wait-list design
Child welfare	1	0.36	1	1	1
Education	1	0.43	0.23	0.22	n/a
Crime	1	0.36	0.80	0.50	n/a
Substance abuse prevention	1	0.33	1	1	1
Substance abuse treatment	1	1	1	1	1
General prevention/public health	Level 1 =0.31 All others = 1	0.38	1	1	1
Early childhood education	1	1	1	1	n/a
Health care (exceptions noted below)	1	1	1	1	1
Asthma self-management education	1	0.36	0.5	1	1
Child depression	1	0.64	1	1	0.44
Child anxiety	1	1	1	1	0.41
Child posttraumatic stress	1	1	1	1	0.36
Child disruptive behavior & ADHD	1	0.56	0.37	1	1
Adult depression and anxiety	1	0.79	1	1	0.46
Adult posttraumatic stress	1	0.63	1	1	0.68
Serious mental illness	1	1	1	1	1

2.4g. Calculating Inverse Variance Weights and Standard Errors when WSIPP Adjustments are made to Effect Sizes.

When we make multiplicative adjustments to effect sizes, we also make adjustments to the standard errors and inverse variance weights. For continuous outcomes, we use equation 2.7 to calculate the adjusted variance (Var_{ad}) substituting the adjusted ES (ES_{adj}) for ES.

For dichotomous outcomes reported as odds ratios or percentages, we first calculate the odds ratio (OR_{adj}) associated with the ES_{adj} using the following equation:

$$(2.36)\ OR_{adj} = e^{(1.65\ ES_{adj})}$$

Next we calculate the corresponding treatment percentage, assuming the comparison rate does not change.
Finally, we calculate the variance per equation 2.10 using the adjusted percentages to estimate values for O_{1t}, O_{2t}, O_{1c}, and O_{2c}.

For dichotomous outcomes reported as chi-square, p-value, or odds ratios and confidence intervals, we first calculate Var_{adj} using equation 2.7 and ES_{adj}. Then, based on our analysis, we multiply the Var_{adj} by 1.65 to provide a good approximation of $Var_{adj}Cox$.

In all cases, the adjusted standard error is the square root of the variance.

2.5 WSIPP Adjustments to Effect Sizes from Longitudinal Linkage Studies

As with the results from program evaluations (discussed in Chapter 2.4), we would ideally make adjustments to the effect sizes from studies measuring the relationship of one outcome to another based on findings from meta-regression. Our current links do not use multipliers, due either too few articles to perform meta-regression or a failure to reject a null hypothesis. The following section describes the procedures we would use if they were available. For any linkage study, we may make up to three types of adjustments that we deem necessary to increase our confidence in the evidence for a causal relationship between two outcomes. We may 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 in Washington; and c) the relevance of the independent and dependent measures that individual studies examined.

2.5a Methodological Quality.

We require a minimum level of methodological quality to be considered in the analysis. To establish that one outcome leads to another, we prefer longitudinal studies that establish clear temporal ordering—where a first outcome (e.g., juvenile crime) precedes another outcome (e.g., high school graduation). Ideally, a study would statistically control for both observable factors and unobservable variables by using fixed effects modeling, natural experiments, twin studies, instrumental variables, or other techniques. Some outcome-on-outcome studies do not have the advantage of longitudinal datasets and they may use cross-sectional data; the results from these studies may be useful, but they may not have as much information to make cause-and-effect inferences.

To track the differences in the quality of research designs for linkage studies, we use a 6-point scale (with values ranging from 0 to 5) as a way to adjust the reported results in a study. On this scale, a rating of 5 reflects a study in which the most confidence can be placed: a longitudinal study with clear temporal ordering and good controls for both observable and unobservable confounds. A rating of 0, on the other hand, reflects a study in which temporal ordering is not established, and we cannot infer a causal link between independent and dependent variables.

On the WSIPP 0-to-5 scale, each linkage study is rated as follows:

5—longitudinal study with temporal ordering and good statistical controls for observable and unobservable confounds

4—longitudinal study with temporal ordering and good statistical controls for observable confounds

3—longitudinal study with temporal ordering but not as many observable controls

2—cross-sectional study with temporal ordering and retrospective measurement of prior outcomes

1—a WSIPP placeholder rating that is not currently used

0—a study for which we cannot infer a causal link between independent and dependent variables

In our meta-analyses, we do not use the results from studies rated as a 0 or 1 on this scale.

Using this scale, if we had a large enough number of studies in a research area, we would conduct a meta-regression to determine if, on average, different research design characteristics affect average effect sizes of the relationship between one outcome and another. Again, our current linked effect sizes do not include multipliers, usually due to too few articles to perform meta-regression.

2.5b Generalizability of the Sample

We may also adjust the effect sizes for linked outcomes for the degree to which the individuals included in the study sample are representative of the Washington population as a whole. If, via meta-regression, we determine that a sample is not representative of the Washington State population, we may use a multiplicative factor to adjust the effect size downward.

2.5c 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 adjust the effect, via a meta-regression analysis. 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. If we had a large enough number of studies, we could then conduct a meta-regression analysis to estimate whether the presumed inferior outcome measures affect, in a systematic manner, the strength of the relationships.

2.6 Meta-Analytic Procedures: Calculating "Adjusted" Effect Sizes for the Benefit-Cost Model

Once all WSIPP adjustments to effect sizes have been made (as described in Chapters 2.4 and 2.5) to the unadjusted effect sizes for each study we review, we then re-run the random effects inverse-variance weighted meta-analysis using equations 2.25 through 2.32, substituting the WSIPP-adjusted effect sizes in lieu of those originally coded from the studies. The results of this second-stage meta-analysis produce the effect size and standard error that we then use in WSIPP's benefit-cost model. At this point in time, we do not calculate adjusted effect sizes for links; as we collect more research evidence, we will attempt to do this in the future.

2.7 The Persistence of Effect Sizes over Time

The benefit-cost model implemented by WSIPP, as illustrated in equation 2.1, anticipates that most programs and policies analyzed will have annual streams of benefits and costs that occur over many years, not just at one point in time. That is, calculating the net present-value of an investment requires information on the long-term changes to annual cash and resource flows. It is important for benefit-cost analysis, therefore, to be able to model effects as they occur over time, judging both when effects occur over the life course, and whether effects change over time.

As we describe in detail in Chapter 3, WSIPP's benefit-cost model explicitly requires two user-supplied time-dimensioned effect sizes. Most often, the research evidence from the meta-analyses will be conducted for outcomes that are observed within the first year or two following program participation. For example, the typical follow-up period for program evaluations of criminal justice programs is about one year. Rather than simply assume that this near-term effect size (and standard error) persists in perpetuity or, on the other hand, drops to zero in year two, the WSIPP model allows the inclusion of a second effect size (and standard error).

We use various procedures to estimate the second effect size (and standard error) depending on the available information. When a topic has enough studies with extended follow-up measurements, our preferred approach is to calculate program specific meta-analyses at various follow-up periods to estimate the second effect size and its standard error. We compute these second effect sizes using steps identical to those described in Sections 2.3 to 2.6.

Unfortunately, many programs do not have enough research to conduct a program-specific meta-analysis to obtain a second effect size. In these cases, we use information from a broader group of research studies that we can apply to any program within that area. We combine effect sizes from all programs in a given research area and regress the effect size on the follow-up period to estimate the relationship between follow-up period and effect size. Depending on the research area and available information, we may use only the longest follow-up from each study or use all follow-up periods from a given study.[35] We test various functional forms and types of models (fixed and random effects, clustered on topic and/or study) within a research area to determine the best model based on overall fit and model interpretation. In a typical meta-regression analysis, we first determine whether follow-up period is a statistically significant predictor of effect size (we use a p-value<0.10 standard); if not, we generally do not adjust our first effect size.

If the effect size does seem to grow or decay over time, we estimate the second effect size in one of two ways:

- We use our preferred regression model or meta-analysis to predict an effect size and standard error at a specific follow-up period; or[36]
- We calculate a multiplicative adjustment (and standard error) from the regression or meta-analysis for a given follow-up period that we apply to a program's first effect size to estimate the second effect size. The second approach may be used if we find that the effect size decays, but we do not suspect that it decays to zero. For example, we may find that, on average, effect sizes decay by 50% over 36 months, but may not decay following those 36 months. For a program for which we have little or no longer-term information, we would multiply the first effect size by 0.5 to get an estimate of the second effect size three years later. We also calculate a standard error on the decay multiplier of 0.5 and use the formula for the variance of the product of two random variables to calculate a standard error for the second effect size.[37]

Finally, in some cases we are unable to estimate program effects beyond the first effect size using either meta-analysis or regression analysis. This typically occurs with "secondary" outcomes. Secondary outcomes are those that are not the prime focus of a program, such as crime outcomes from studies whose primary focus is changes in substance abuse outcomes. In these cases, we often have few or no rigorous evaluations that measure the outcome over time and thus we cannot predict whether program effects on these secondary outcome decay over time. For these secondary outcomes, until more information is accumulated, we assume that effects decay to zero for all time periods following those measured in the studies.

[35] When including multiple follow-up periods from a given study, we cluster our standard errors by study.
[36] We typically carry out the prediction in STATA with the lincom command.
[37] We typically predict the multiplier and the standard error with STATA's nlcom command.

Exhibit 2.1
Current WSIPP Decay Factors by Outcome

Outcome	ES at time 2	SE at time 2	Time 2
Child abuse & neglect	ES1	SE1	Age 17
Out-of-home placement	ES1	SE1	Age 17
Substance abuse prevention outcomes	ES1	SE1	Age at Time 1 + 10
Substance abuse treatment outcomes For most programs Contingency management (higher-cost) Contingency management (lower-cost)	 0 0 0	 0.187 0.125 0.075	 Age at Time 1 + 3 Age at Time 1 + 1 Age at Time 1 + 1
Substance abuse outcomes Brief intervention strategies	ES1 * 0.137	$\sqrt{(SE1^2 * 2.25)}$	Age at Time 1 + 2
Crime	ES1	SE1	Age at Time 1 + 10
Adult depression, adult anxiety	ES1 * 0.52	$(SE1^2 * 1.5)^{0.5}$	Age at Time 1 + 2
Adult PTSD	ES1	SE1	Age at Time 1 + 1
Adult psychosis	ES1 * 0.743	$(ES1^2 * 0.569^2 + 0.743^2 * SE1^2 + SE1^2 * 0.569^2)^{0.5}$	Age at Time 1 + 1
Child PTSD	ES1	SE1	Age at Time 1 + 1
Child ADHD	0	$(ES1^2 * 0.048^2 + 0.00317^2 * SE1^2 + SE1^2 * 0.048^2)^{0.5}$	Age at Time 1 + 1
Child depression	ES1 *0.00099	$(ES1^2 * 0.0811^2 + 0.00099^2 * SE1^2 + SE1^2 * 0.0811^2)^{0.5}$	Age at Time 1 + 1
Child anxiety	ES1*0.4623	$(ES1^2 * 0.0992^2 + 0.4623^2 * SE1^2 + SE1^2 * 0.0992^2)^{0.5}$	Age at Time 1 + 1
Child internalizing	ES1 * 0.72848	$(ES1^2 * 0.2803^2 + 0.7285^2 * SE1^2 + SE1^2 * 0.2803^2)^{0.5}$	Age at Time 1 + 2
Child externalizing, child disruptive behavior	ES1 *0.47646	$(ES1^2 * 0.2012^2 + 0.47646^2 * SE1^2 + SE1^2 * 0.2012^2)^{0.5}$	Age at Time 1 + 3
Psychiatric hospitalization Assertive community treatment ER prevention for frequent users	0	0.118	Age at Time 1 + 1
Diabetes	ES1 * 0.478	0.077	Age at Time 1 + 7
Weight change Intensive/long-term diabetes interventions Short-term diabetes interventions Obesity prevention for children Obesity prevention, adults, high-intensity Obesity prevention, adults, low-intensity	 0 ES1 * 0.31 0 0 0	 0.054 0.101 0.07 0.012 0.012	 Age at Time 1 + 7 Age at Time 1 + 7 Age at Time 1 + 2 Age at Time 1 + 5 Age at Time 1 + 2
Obesity Obesity prevention for children Obesity prevention, adults, high-intensity Obesity prevention, adults, low-intensity	 0 0 0	 0.101 0.086 0.086	 Age at Time 1 + 2 Age at Time 1 + 5 Age at Time 1 + 2
Emergency room visits for asthmatic children or general population	0	0.0861	Age at Time 1 + 2
Hospitalizations (readmissions) PCMH Outcomes for seriously mentally ill individuals, those easily lost to follow up	0	0	Age at Time 1 + 1
Labor market earnings (measured directly) Case management programs Job search and placement Training, no work experience Training with work experience Work experience	 0 0 0 0 0	 0.014 0.017 0.032 0.018 0.0013	 Age at Time 1 + 1 Age at Time 1 + 2 Age at Time 1 + 1 Age at Time 1 + 1 Age at Time 1 + 2

Chapter 3: Procedures to Compute "Monetizable" Outcome Units from Effect Sizes

Chapter 2 described the procedures WSIPP uses to compute effect sizes and standard errors from meta-analyses. This Chapter describes our procedures to convert effect sizes into units of outcomes that can be monetized. Chapter 4 then describes how monetary values are attached to these "monetizable" outcome units.

The procedures in this chapter are necessary because WSIPP's model uses "program effect sizes" rather than simply "program effects." This seemingly arcane distinction is important for our approach to benefit-cost modeling.

- A **"Program Effect."** A finding from an individual program evaluation produces an estimate of whether the program had an effect on an outcome. For example, a K–12 tutoring program may improve high school graduation rates by 4 percentage points—from, say, 75% without the program to 79% with the program. This is a program effect. An effect—in this example, a four percentage point gain in the probability of high school graduation—can be monetized directly with the procedures we describe in Chapter 4. If we were only interested in conducting a benefit-cost analysis based on the finding of a single program evaluation, we would not need the procedures we describe in Chapters 2 and 3. Rather, we would simply observe the percentage point change and proceed directly to Chapter 4 to monetize the program effect.
- A **"Program Effect Size."** WSIPP, however, desires to draw an overall conclusion about a topic by considering all credible research studies on the topic, not just the results of a single study. Because of this, for each program evaluation we review, we first convert a program effect into an effect size metric, with the procedures described in Chapter 2. With this common metric, we are then able to meta-analyze a collection of studies on a single topic. While this process gains us all of the advantages that come from conducting a meta-analysis, the downside is that in order to perform a benefit-cost analysis we must re-convert the meta-analyzed effect size back into a program effect—measured in the natural units of the particular outcome. In other words, a meta-analyzed effect size cannot be directly monetized by itself; it must first be re-converted into a program effect.
- A **"Program Unit Change."** For purposes of clarity in this presentation, we call a program effect a "unit change" in order to clearly separate the concept from that of an effect size. This Chapter describes how we compute unit changes from the effect sizes we describe in Chapter 2.

To continue the K–12 tutoring example above, we would compute a D-cox effect size, using equation 2.9, of +0.137 for the four percentage point program effect in the hypothetical program evaluation. We would then make similar effect size calculations for all of the tutoring studies in our meta-analysis and might conclude, for example, that the weight of the evidence finds that tutoring programs, on average, can be expected to have a D-cox effect size of +0.15 on high school graduation. From this effect size finding, in order to compute a metric that can be used in benefit-cost analysis, we would apply the procedures described in this Chapter to compute a unit change for the tutoring topic.

Not all program effect sizes are used in the final benefit-cost calculation. For example, we are currently unable to translate some effect sizes into monetizable units, but we report the effect size as the outcome is still of interest to legislators and other audiences. Some effect sizes trigger the same monetization routines as other effect sizes in a meta-analysis. When this happens, the monetizable units are compared against each other, and one effect size may "trump" another in the same analysis (see Chapter 5 for a detailed discussion of these procedures). Additionally, in some instances where there is only one study or a limited or non-representative sample, WSIPP may only report the program effect sizes from the meta-analysis. Finally, if the research literature has several outcomes measured in multiple studies and some that are measured in only one study with a limited or non-representative sample, WSIPP may choose in include only the outcomes with multiple effect sizes in the benefit-cost analysis.

3.1 Effect Size Parameters from Program Evaluations

As noted in Chapter 2, the WSIPP benefit-cost model monetizes changes to outcomes measured as quantities. For example, outcome quantities might be crimes avoided, increases in high school graduation rates, increases in student standardized test scores, or reductions in the probability of child abuse and neglect. Depending on whether these outcome quantities are measured as dichotomies or on continuous scales, the general information needed to compute quantities includes an effect size (ES) and certain *Base* information about the population being served by a program. This is given in the following equation:

$$(3.1)\ Q_y = f(ES, Base)$$

In the WSIPP benefit-cost model, equation 3.1 is operationalized with several user-supplied parameters. For each topic for which a benefit-cost analysis is to be calculated, these parameters include the following:

Tage	average age of a person treated with a program
Mage1	average age of a person when the first effect size for a particular outcome of the program is measured
ES1	estimated effect size for a particular outcome of a program at *Mage1*
$ES_{SE}1$	estimated standard error of the effect size for a particular outcome of a program at *Mage1*
Mage2	average age of a person when a second effect size for a particular outcome of the program is measured
ES2	estimated effect size for a particular outcome of a program at *Mage2*
$ES_{SE}2$	estimated standard error of the effect size for a particular outcome of a program at *Mage2*
Base	estimated outcome for the non-treatment group (e.g., the 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. The *Base* may change with the age of the participant; it is not necessarily a single number. In many cases, the *Base* increases year-on-year, representing, for example, the cumulative likelihood of criminal activity over time, or the cumulative likelihood of child abuse or neglect over time.

The user enters the age of the person treated when the first program effect for a particular outcome was measured; we call this *Mage1*. If the user has conducted a meta-analysis, this first measurement age should represent the average follow-up period in the underlying program evaluations in the meta-analysis. For example, in juvenile justice literature, criminal recidivism typically is measured one or two years following treatment. The user will also enter the other two parameters centered on this first measurement age: the effect size, *ES1*, and its standard error, $ES_{SE}1$, as calculated with the procedures in Chapter 2.

Next, the user repeats this sequence for a second measurement period for a particular outcome. That is, a user enters the age of the person treated when a second program effect was measured or projected; we call this *Mage2*. *Mage2* will always be greater than *Mage1*; it is designed as a way to project the longer run effectiveness of a program. Program effects could decay, grow, or stay the same as time passes, following treatment. The second follow-up period allows the modeling of the trajectory of these longer–run effects. The user will also enter the other two parameters centered on this second measurement age: the effect size, *ES2*, and its standard error, $ES_{SE}2$.

Many program evaluations do not measure effect sizes at multiple follow-up periods. Therefore, it is unlikely that the second period effect sizes will come from the procedures described in Chapter 2. If, however, the user has conducted a meta-regression, it may be possible to make inferences about the longer run effect sizes. As noted in Chapter 2.7, increasingly WSIPP conducts meta-regressions to inform our projection of longer-term program effect sizes.

For example, in the juvenile justice program called Functional Family Therapy (FFT), the assumed treatment age for the average juvenile in this program is 17. Next, the user inputs six of the eight parameters for the crime outcome measured for FFT. The first effect size is -0.261 and has a standard error of 0.096. For this program, our review of the FFT evaluations indicates that the average follow-up period is about two years; thus, we enter age 19 as *Mage1*. The second effect size, -0.261, is entered for age 29 with a standard error of 0.096. In the case of juvenile justice programs, the longer-term outcome is the same as that entered at the first follow-up period because our meta-regressions have indicated that effects of programs on crime effects do not appear to fade out as time passes. In outcomes in other public policy areas, K–12 student test scores for example, we have found through meta-regressions that test scores effects decay over time. The WSIPP model accommodates the modeling of these time-dimensioned outcomes with this two point process.

The user selects the appropriate population for each outcome affected by a program. The actual *Base* rates for each program outcome are input separately within the model. For example, for education outcomes, the user selects whether a program affects all students or low-income populations. This selection will then direct the model to use the base inputs (high school graduations rates, test score information, and other parameters) entered elsewhere in the model.

3.2 Monetizable Unit Changes from Effect Sizes from Program Evaluations

Once these eight parameters are exogenously computed and entered into the model software, we follow several steps to compute monetizable "unit changes." We begin by computing unit changes for each outcome directly measured by the program evaluations. The unit changes are the quantity of change in outcomes we can expect from a program or policy, compared to the outcomes of people who do not receive the program.

For continuously measured outcomes, as given by equations 3.2 and 3.3, the change in units at the first and second measurement ages, *Mage1* and *Mage2*, is calculated simply with a Cohen's *d* effect size and a *Base* variable, which is measured as a standard deviation of the outcome measurement.

$$(3.2)\ \ Q_{Mage1} = Base_1\ \times ES_1$$

$$(3.3)\ \ Q_{Mage2} = Base_2\ \times ES_2$$

1. We distribute the unit change calculated at *Mage1* (equation 3.2) to the ages between *Tage* and *Mage1*.
2. We distribute the unit change calculated at *Mage2* (equation 3.3) to ages *Mage2* and after.
3. For ages ranging from *Mage1* to *Mage2*, we linearly interpolate the unit change between *Mage1* and *Mage2*.

In Monte Carlo simulations, equations 3.2 and 3.3 are implemented using random draws from a normal probability density distribution with the effect size (ES_1 and ES_2) and its standard error (ES_{se1} and ES_{se2}). A common randomly drawn seed is used to compute both Q_{mage1} and Q_{mage2} for each Monte Carlo case.

For dichotomously measured outcomes, as given by equations 3.4 and 3.5, the change in units (percentage point changes in the outcome), Q_{mage}, at the first and second measurement ages, *Mage1* and *Mage2*, is calculated with a D-cox effect size and a *Base* variable, which is measured as a percentage. Exhibit 3.0 provides a numeric example to illustrate these procedures for dichotomous outcomes, which is slightly more involved than that for continuous outcomes.

$$(3.4)\ \ Q_{Mage1} = \left(\frac{(e^{ES_1 \times 1.65} \times Base_1)}{(1 - Base_1 + Base_1 \times e^{ES_1 \times 1.65})} - Base_1\right)$$

$$(3.5)\ \ Q_{Mage2} = \left(\frac{(e^{ES_2 \times 1.65} \times Base_2)}{(1 - Base_2 + Base_2 \times e^{ES_2 \times 1.65})} - Base_2\right)$$

$$(3.6)\ \ Q_{seMage1} = ABS\big(Q_{Mage1} \times ES_{se}1/ES1\big)$$

$$(3.7)\ \ Q_{seMage2} = ABS\big(Q_{Mage2} \times ES_{se}2/ES2\big)$$

1. Equations 3.4 and 3.5 compute the percentage change in a dichotomous outcome (Q_{Mage1} and Q_{Mage2}) measured at the two ages, *Mage1* and *Mage2*, using the D-cox effect size formula (see Chapter 2). The unit change is calculated with the effect sizes at the two ages and is calibrated relative to the base rate for the outcome measured at *Mage1* and *Mage2*, respectively. In the example calculation in Exhibit 3.0 we show this in columns (2), (3), (5), and (6).
2. The standard errors ($Q_{seMage1}$ and $Q_{seMage2}$) of the unit changes at *Mage1* and *Mage2* are calculated using equations 3.6 and 3.7. The standard errors are the absolute value of the product of the unit change (Q_{mage}) times the coefficient of variation (ES_{se} / ES) in the effect sizes at each age. In the example calculation in Exhibit 3.0 we show this in columns (3), (10), and (11).
3. For ages ranging from *Tage* to *Mage1*, we distribute the percentage change calculated at *Mage1* to the ages between *Tage* and *Mage1* and then multiply the percentage change by the base rate at each age. In the example calculation below, we show this in columns (8) and (9).
4. For ages beyond *Mage2*, we distribute the percentage change calculated at *Mage2* to ages *Mage2* and after and then multiply the percentage change by the base rate at each age. In the example calculation below, we show this in columns (8) and (9).

5. For ages ranging from *Mage1* to *Mage2*, we linearly interpolate the percentage change between *Mage1* and *Mage2* and then multiply the percentage change by the base rate at each age. In the example calculation below, we show this in columns (8) and (9).
6. For the standard errors in the unit changes for ages ranging from *Tage* to *Mage2*, we distribute the coefficient of variation calculated at *Mage1* and then multiply the coefficient by the unit change at each age. In the example calculation below, we show this in columns (10) and (11).
7. For the standard errors in the unit changes for ages from *Mage2* and beyond, we distribute the coefficient of variation calculated at *Mage2* and then multiply the coefficient by the unit change at each age. In the example calculation below, we show this in columns (10) and (11).
8. When the model is run in Monte Carlo mode, the unit change is calculated for each year with a normal probability density distribution with a mean (column (9) in the example) and the standard error (column (11) in the example). A common random seed is used for all years for each draw of a Monte Carlo simulation. For these dichotomous outcomes, it is possible with the procedures above to draw an outcome above 1.0 or below 0. To avoid this illogical draw, we implement bounding rules. If a random draw results in a unit change that produces an outcome above 1.0 probability, then that draw is set so that the unit change produces an outcome probability equal to 1.0. Similarly, if a random draw results in a unit change that produces an outcome below zero probability, then that draw is set so that the unit change produces an outcome probability equal to zero.

Exhibit 3.0

Example of Procedure for Computation of Dichotomous Outcome Unit Changes

	Load the Exogenous Information			Compute Changes at Mage1 and Mage2			Compute Unit Changes and Standard Errors for All Years			
age	Load the two effect sizes at Mage1 and Mage2	Compute the coefficient of variation at mage1 and mage2	Load base rates for the outcome at each follow up age	Compute the treatment group rate at mage1 and mage2	Compute the unit change (pct points)	Compute the percentage change	Distribute the percentage change to other years	Compute Unit Change (pct points)	Distribute the coefficient of variation	compute the standard error on the Unit Change
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
34			0.240				-0.229	-5.5%	0.500	0.027
35	-0.200	0.500	0.240	0.185	-0.055	-0.229	-0.229	-5.5%	0.500	0.027
36			0.240				-0.083	-2.0%	0.500	0.010
37	0.050	1.000	0.240	0.255	0.015	0.064	0.064	1.5%	1.000	0.015
38			0.240				0.064	1.5%	1.000	0.015
39			0.240				0.064	1.5%	1.000	0.015
40			0.240				0.064	1.5%	1.000	0.015
41			0.240				0.064	1.5%	1.000	0.015
42			0.240				0.064	1.5%	1.000	0.015
43			0.240				0.064	1.5%	1.000	0.015
44			0.240				0.064	1.5%	1.000	0.015
Inputs										
34	Tage (age of person at time of treatment)									
35	Mage1 (age of person when outcome first measured)									
-0.200	ES1 (effect size at Mage1)									
0.100	SE1 (standard error at Mage1)									
37	Mage2 (age of person when outcome is measured a second time)									
0.0500	ES2 (effect size at Mage2)									
0.050	SE1 (standard error at Mage1)									

3.3 Linked Effect Size Parameters

As noted in Chapter 2.1, one of the characteristics of WSIPP's approach to benefit-cost modeling is the inclusion of research that establishes how one outcome is linked to another outcome. In expression (3.8), these linkages are the relationships between $Outcome_1$ and $Outcome_2$.

$$(3.8) \quad if\ Program \rightarrow Outcome_1, \qquad and\ if\ Outcome_1 \rightarrow Outcome_2, \quad then\ Program \rightarrow Outcome_2$$

The benefit-cost model then uses these linkages to supplement the direct findings from program evaluations (shown in the expression as the direct effect of a $Program$ on $Outcome_1$). The magnitude of these linkages are estimated with the meta-analytic procedures describe in Chapter 2, although we do not measure or predict an effect size at a second time period (or decay factor). The linkages are computed with the estimated mean effect size and standard error of relationships between outcomes measured in evaluation studies, and other monetizable outcomes. For example, crime as a juvenile reduces the probability of high school graduation (and the resulting labor market earnings boost that high school graduation allows). Crime has an effect size of -0.393 on earnings via high school graduation, with a standard error of 0.091. The "age at which relationship begins" is indicated as 18; this means that the monetary benefits of linked high school graduation through crime begin at age 18. This also means that if a program has a direct impact on the crime after age 18, then it is too late to activate these linked benefits of high school graduation. For links that do not occur at a specific, consistent point in time (such as the effect of alcohol use in middle school on future alcohol use disorder), we apply the linked effect to all years following program intervention. We list our current estimates for the linkages in this report's Appendix.

3.4 Unit Changes from Linked Effect Sizes

For linkages between outcomes the user enters a single effect size, standard error, and age of the person to whom the measurement applies. To compute the linked unit change from these link effect sizes, we follow analogous procedures to those described in Chapter 3.2, above.

For continuous outcomes, as shown in equation 3.9, the linked unit change at each age is simply the linked effect size at $LinkAge$, multiplied by the standard deviation unit in which the outcome is measured using the following equation:

$$(3.9) \quad LinkQ_{LinkAge} = Base \quad \times ES_{link}$$

For dichotomous outcomes, as shown in equation 3.10, the linked unit change for linked effect sizes is computed as described in the previous section. We first compute the percentage change in the outcome measured for the linked effect size at the age of the link supplied by the user, using the D-cox effect size formula (see Chapter 2).

$$(3.10) \quad LinkQ_{LinkAge} = \left(\frac{(e^{ES_{link} \times 1.65} \times Base)}{(1 - Base + Base \times e^{ES_{link} \times 1.65})} - Base \right)$$

3.5 Monetizable Unit Changes for Benefit-Cost Calculation, When a Linked Outcome is Present

When a linked outcome has been established and entered, the model will use the result to complete the steps in expression 3.8 (on the previous page). As the model runs, it searches for any possible links to the direct program outcomes measured, and then implements the procedures in Chapter 3.3 and 3.4. The linked unit of change (*Program* on $Outcome_2$) is simply the multiplicative product of the unit change from the program evaluation (*Program* on $Outcome_1$) and the unit change from a relevant link ($Outcome_1$ on $Outcome_2$). We do not currently estimate links from outcomes measured with elasticities or semi-elasticities.

To illustrate the computations with hypothetical numbers, suppose that the juvenile justice program Functional Family Therapy (FFT) reduces a juvenile's crime probability by 10 percentage points. This is the program unit change as described in Chapter 3.2 (*Program* on $Outcome_1$). Further, suppose that a juvenile that engages in crime has a reduced probability of graduating from high school of 20 percentage points. This is the linked unit change as described in Chapter 3.4 ($Outcome_1$ on $Outcome_2$). Then, multiplying these two changes, FFT can be expected to increase the high school graduation probability (*Program* on $Outcome_2$) by 2 percentage points (0.10 X 0.20 = 0.02). That is, if the evaluations of FFT had measured high school graduation as an outcome, we would have expected the result to have been a 2 percentage point increase in high school graduation probability. When the benefit-cost model is run, Monte Carlo simulation is used to estimate this linked relationship and its standard error (with random draws from normally distributed mean effects and standard errors for the first two steps in the expression). In the benefit-cost model, the benefits of FFT will then be compute for a 10% change in crime outcomes and a 2% change in high school graduation. Again, these particular numbers are hypothetical and for illustration purposes only; our actual current estimates for FFT are different than this illustrative example.

Chapter 4: Procedures to Estimate Monetary Benefits of Outcome Units

As summarized in Chapter 1, the WSIPP 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 the following equation:

$$(4.1)\quad NPV_{tage} = \sum_{y=tage}^{N} \frac{Q_y \times P_y - C_y}{(1+Dis)^y}$$

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, multiplied by 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 the annual cash flows 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 4.1 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 4.1, 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 Chapters 2 and 3. In this Chapter, we describe the various methods we use to estimate the price term, P_y, in equation 4.1.

4.1 Valuation of Labor Market Outcomes

Several of the outcomes measured in the benefit-cost model are monetized with how a program-induced change in an outcome affects lifetime labor market earnings. Measuring the earnings implications of human capital variables is a common approach in economics.[38] Chapter 4.1a discusses the common data sources we use for all of the estimates involving labor market earnings. Other parts of Chapter 4 present additional outcome-specific parameters, along with the computational routines, to produce estimates of labor market earnings.

In the current version of the benefit-cost model, the following outcomes are monetized, in part, with how changes in an outcome affect labor market earnings (see Chapter sections in parentheses for more information on each outcome):

- High school graduation (Chapter 4.7)
- Standardized student test scores (Chapter 4.7)
- Number of years of completed education (Chapter 4.7)
- Morbidity and mortality costs of alcohol and illicit drug disorders, and regular smoking (Chapter 4.4)
- Morbidity and mortality costs of mental health disorders (Chapter 4.8)
- Morbidity and mortality costs of health care outcomes (Chapter 4.9)
- Morbidity and mortality costs of child abuse and neglect (Chapter 4.3)

One way the model organizes earnings is by educational subgroup. These educational subgroup calculations are described in Section 4.1b. In addition, the benefit-cost model estimates earnings streams and employment rates by populations relevant to the workforce at large. These calculations are described in Section 4.1c. Calculations of variations in labor market earnings and employment by various health conditions, mental health disorders, and substance use disorders are described in Section 4.1d. The valuation of household production is described in Section 4.1e. Finally, outcomes may directly change earnings or change earnings through the probability of employment. These calculations are described in Section 4.10.

[38] See, for example, Heckman, J.J., Humphries, J.E.,& Veramendi, G. (2015). *The Causal Effects of Education on Earnings and Health*, Working Paper March 12, 2015. See also, Rouse, C.E. (2007). Consequences for the labor market. In Belfield, C.R. & Levin, H.M. (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). *Some simple analytics of school quality* (NBER Working Paper No. 10229). Cambridge, MA: National Bureau of Economic Research.

4.1a Calculating Earnings

Earnings Data and Related Parameters. In the benefit-cost model, all earnings related estimates derive from a common dataset. The estimates are taken from the outgoing rotation of the US 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.[39] To keep the model as simple as possible, we gather "person variables" from the CPS summary files: 1) PEARNVAL, person total earnings—this variable measures income from earnings, not total money income and 2) A_AGE, age by single year. These data are representative of the US population, not just those living in Washington State.

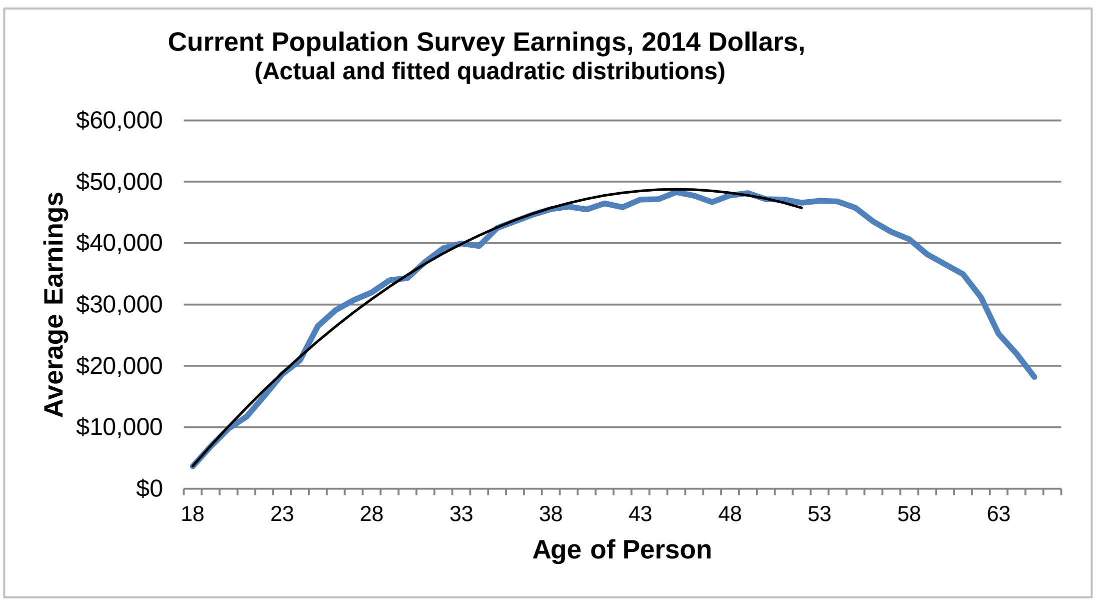
Due to the cyclical nature of the business cycle, we use data that attempts to match the November 2001 to June 2009 business cycle as reported by NBER.[40] We use the 2002 through 2010 March CPS files, as the March CPS covers earnings for the prior year. The sample was restricted to persons age 18 to 65 inclusive and weighted by the CPS march supplement final weight scaled such that the sum of the weights is equal to the number of unweighted observations in the data. From this sample, we ran a regression to compute average earnings per person by single year of age. We refer to this as *CPSEarnAll*.

This regression was run in SAS (9.4) using PROC REG as given by the following equation:

$$(4.2)\quad EarnAll_y = \beta 0 + \beta 1 * \mathrm{AGE} + \beta 2 * \mathrm{AGE}^2 + \beta 3 * \mathrm{YR}_{2003} + \beta 4 * \mathrm{YR}_{2004} + \ldots + \beta 10 * \mathrm{YR}_{2010}$$

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 CPS data series we include in the model measure both earnings of the earners and the rate of labor force participation. This distinction becomes important when we discuss how these earnings estimates are used to monetize specific outcomes. The raw CPS earnings data and the fitted curve from the predicted values of the regression are plotted below. Numbers are inflated to 2014 dollars using the IPD described in more detail in Section 4.11f. Futher adjustments, described below, adjust the data to match the future labor market in Washington.

Exhibit 4.1



[39] The data are accessed from the "DataFerrett" application of the US Department of Commerce, Bureau of the Census, available from http://dataferrett.census.gov

[40] A business cycle is the length of time between peaks (times when the economy begins to shrink after growing) or between troughs (times when the economy begins to grow after shrinking). The Business Cycle Dating Committee of the National Bureau of Economic Research reports peaks and troughs on its website at http://nber.org/cycles/cyclesmain.html.

State-specific Adjustment for Wages. We use an adjustment ratio to approximate earnings in Washington State relative to the national average in the CPS. We replicate the same procedures used to estimate the earnings by age, including a Washington State dummy variable to measure the differential in earnings between Washington and the rest of the country. For this model, the percent of earnings attributable to the Washington State dummy in the new regressions is used as the adjustment factor.

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 change in earnings. We collect the same cross-sectional CPS information for the last six business cycles—1971 (with data for 1970) to 2010 (with data for 2009).[41] We adjust the series for inflation using the US Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce (see Section 4.11f). We then fit a log-linear model: $\ln(earnings) = a + b(year)$. We correct for autocorrelation with the SAS Proc AutoReg autoregressive model with two lags. We use the coefficients from the model as our real growth rate in earnings.

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 Costs for Employee Compensation (ECEC), which is calculated from the National Compensation Survey (NCS).[42] The ECEC includes paid leave, supplemental pay, insurance, retirement and savings, and legally retired benefits.[43]

Exhibit 4.2

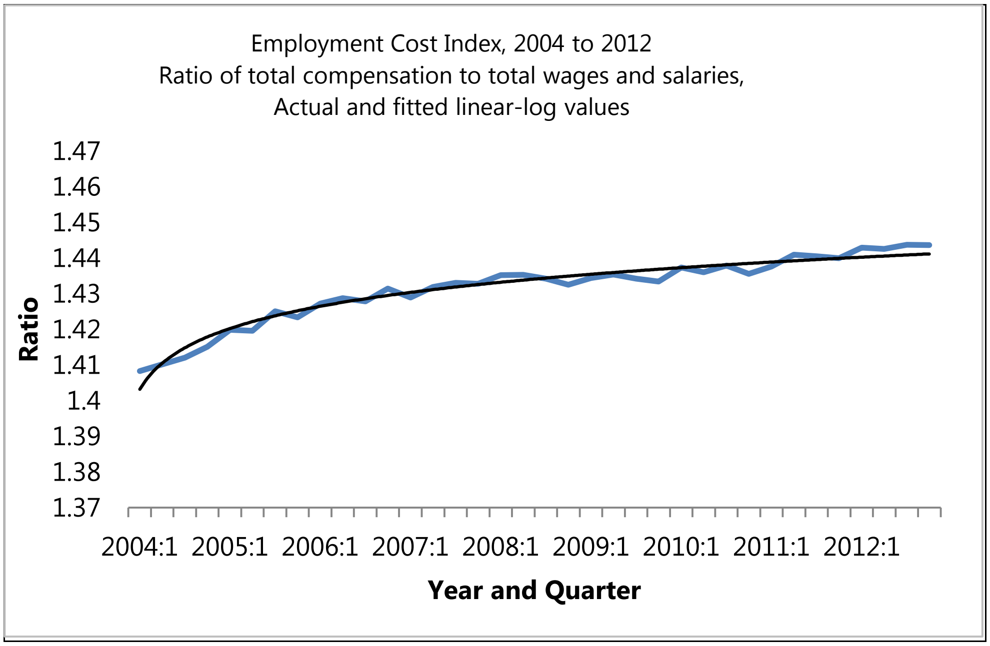
Earnings Adjustment Parameters, General Population

Parameter	Value
Annual real growth rates in earnings	0.0137
Benefits-to-earnings ratios	1.4410
Annual growth rate in the benefits-to-earnings ratio	0.00041
Ratio of State to National median earnings	1.036

[41]We use a sample including persons ages 18-75 for our calculations of the adjustment of Washington State-specific wages and the growth in earnings.
[42]US Bureau of Labor Statistics. (2016). *Employer costs for employee compensation—December 2015* (USDL-16-0463), Washington DC: Author. Data retrieved March 30, 2016 from http://www.bls.gov/news.release/pdf/ecec.pdf
[43]Ibid.

Exhibit 4.3 displays the quarterly national ECEC ratio of total compensation to total wages for all civilian workers. We fit a linear-log model *(ratio = a +b(ln(quarter)))* to the historical series and then forecast the annual values for 2012 and 2042 from which we compute a forecast of the annual rate growth in the benefit ratio over the 30 year interval. The 2014 year benefit ratio and the calculated growth rate are then entered into the model.

Exhibit 4.3



General Mortality Adjustment to Earnings. Within our monetization routines, the change in earnings is estimated by comparing the predicted lifetime earnings of a person who experienced a program with the predicted lifetime earnings of a person who did not. We use CPS data to represent the predicted earnings of that non-participating person. However, the CPS surveys living people, so the numbers do not include the chance that a person has died. Using the general life table described in Section 4.11c, we adjust the predicted labor market earnings for the probability of survival in each year after participation in a specific program or intervention.

The earnings series is then used in the benefit-cost model to estimate labor market-related benefits of a number of outcomes, as described in other sections of this chapter. For example, in each year (y), the basic CPS earnings series is adjusted with the factors described above as given by the following equation:

$$(4.3)\quad ModEarnAll_y = ((EarnAll_y \times (1 + EscAll)^{y-tage}) \times (FAll \times (1 + EscFAll)^{y-tage}) \times (IPD_{base}/IPD_{cps}) \times StateAdjAll\ ) \times ProbLife_y$$

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 for all people ($EscAll$) 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 for all people ($EscFAll$) raised to the number of years after program participation, multiplied by 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}), multiplied by the ratio of state-to-national earnings for all people ($StateAdjAll$), multiplied by the general probability that the person is alive ($ProbLife_y$) to realize those benefits.

4.1b Earnings by Educational Attainment

In addition to the general population, the WSIPP model monetizes the differences in earnings for people of different educational levels to calculate the value of educational attainment (see Section 4.7c). We use the CPS variable A_HGA, educational attainment by the highest level completed to subset the sample by education. We perform the calculations described in Section 4.1a using subsets of the data sample for four educational status groupings:

- 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 no BA degree (*CPSEarnSomeCol*)
- Those with a BA degree or higher (*CPSColDeg*)

For each of these four groups, we replicate the regressions and modeling to determine separate earnings by age distributions and different earnings growth parameters, displayed in Exhibits 4.4 and Exhibit 4.5. We assume that students do not earn money for the time that is spent in higher education, and so for the "some college" and "BA or higher" populations, we set earnings to zero for the expected time spent in college (described in Section 4.7c).

The current BLS data for the ECEC does not allow the index to be broken out by education achievement level. Therefore we enter the same values for benefits for each educational group. It is, of course, likely 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 when relevant inputs can be located.

Exhibit 4.4

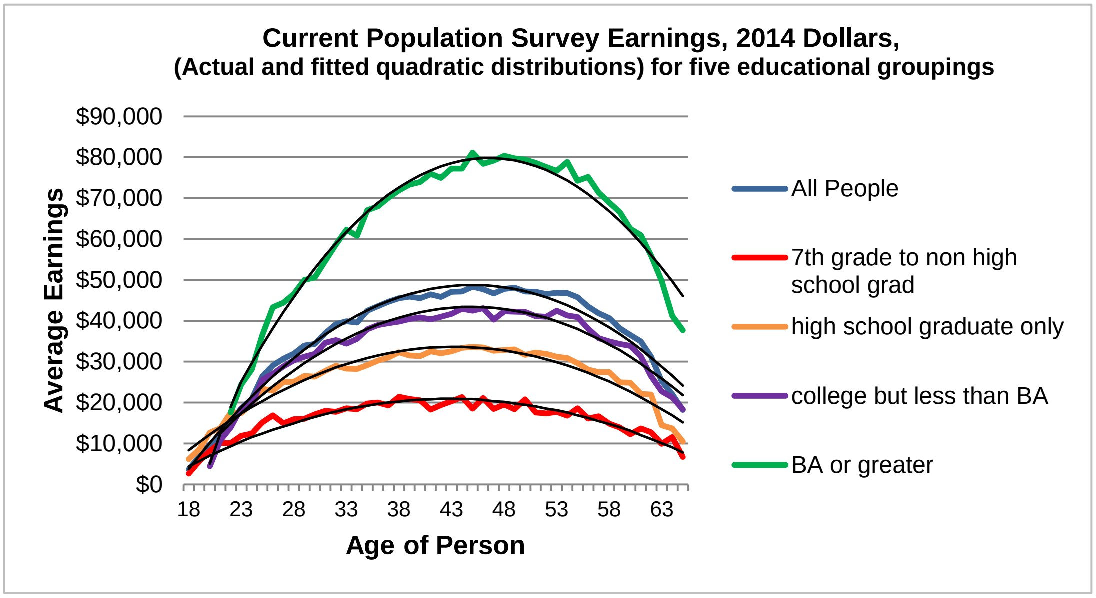


Exhibit 4.5

Earnings Adjustment Parameters by Educational Attainment

	7th grade to non high school	High school graduate only	College but less than BA	BA or greater
Annual real growth rates in earnings	-0.0062	0.0053	0.0095	0.0115
Benefits-to-earnings ratio	1.441	1.441	1.441	1.441
Annual growth rate in the benefits-to-earnings ratio	0.00041	0.00041	0.00041	0.00041
Ratio of state to national earnings	1.079	1.074	1.003	0.935

4.1c Earnings by Population Characteristics

The WSIPP model also values earnings for populations not defined by educational attainment. For example, WSIPP estimates values for some programs that directly target the labor market. We therefore segmented the earnings data into sub-populations that closely align with individuals who participate in different types of workforce training programs. To create these populations we use the following variables from the March CPS supplement data dictionary: A_WKSLK, A_LFSR, A_FAMREL, A_MARITL, and A_HGA. We calculate earnings by age using the methods described in Section 4.1a for four workforce subgroups in addition to that for all people:

- Short-term unemployed (nine or fewer weeks)
- Long-term unemployed (more than nine weeks), non-college graduates
- Not employed single parents
- Not employed single parents (HS education or less)

The calculation of earnings escalation and the state specific adjustment are calculated as the average of the applicable calculated earnings by education subgroups. For each of these four groups, we replicate the regressions and modeling to determine separate earnings by age distributions and to calculate the percent of the subgroup that is employed (has earnings greater than 0). We calculate growth parameters and state adjustment factors based on combinations of relevant education subgroups. Our factors are displayed in Exhibits 4.6 and Exhibit 4.7.

Exhibit 4.6

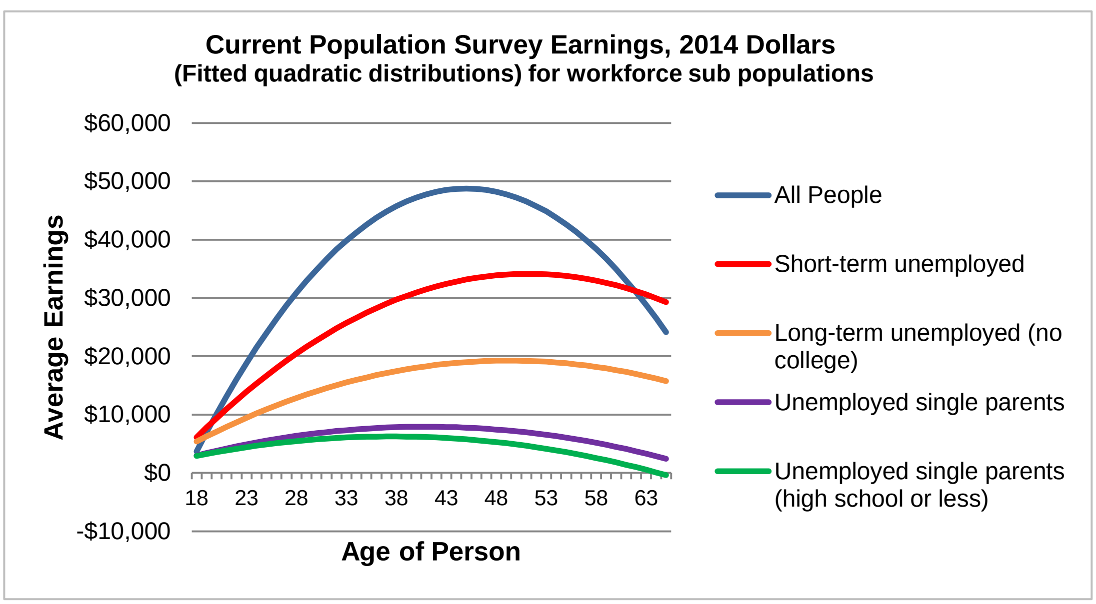


Exhibit 4.7

Earnings Adjustment Parameters by Workforce Population

	All people	Short-term unemployed[1]	Long-term unemployed (no college)[2]	Unemployed single parents[1]	Unemployed single parents (high school or less)[3]
Annual real growth rates in earnings	0.0137	0.0137	0.0028	0.0137	-0.0005
Benefits-to-earnings ratio	1.441	1.441	1.441	1.441	1.441
Annual growth rate in the benefits-to-earnings ratio	0.00041	0.00041	0.00041	0.00041	0.00041
Ratio of state to national earnings	1.036	1.036	1.052	1.036	1.076
Probability of employment	0.770	0.823	0.679	0.391	0.366

[1] Subset of all people.
[2] Average of factors for less than high school, high school graduate, and some college education subgroups.
[3] Average of factors for less than high school and high school graduate education subgroups.

4.1d Earnings and Employment used in Modeling Disease and Disorder

The literature concerning the effects of health conditions, mental health disorders, and substance use on labor market earnings predominantly focuses either on the change in employment status or the change in earnings given employment. The standard analysis of earnings described in the sections above uses a single number for the average earnings of all people whether employed or unemployed. When valuing the changes in labor market earnings due to health conditions, mental health, or substance use disorders, we use the general population from the CPS to estimate base parameters (see Exhibit 4.8). We then apply the effect of the condition or disorder on rate of employment, and the effects of the condition or disorder on the level of earnings if employed (compared to the general population). The procedures we use to compute the value of earnings for various conditions and disorders are described in detail in Section 4.4d.

Exhibit 4.8

Base Assumptions for Earnings and Employment, 2013 CPS

Mean earnings of workers	SD of earnings of workers	Percent of population that works
50,383	64,518	77.7%

4.1e 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 our model for all of the conditions that have a chance of leading to death (i.e., ATOD, mental health disorders, health conditions, child abuse & neglect). 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.[44] 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 hour per day for food preparation, and 2.0 hours per week for household maintenance. These estimates are similar to the 21.4 hours per week calculated by Douglass et al.[45] 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 service.[46]

---

[44] 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

[45] Douglass, J., Kenney, G., & Miller, T. (1990). Which estimates of household production are best? *Journal of Forensic Economics, 4*(1), 25-45.

[46] 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

Exhibit 4.9

Household Production Parameters

Parameter	Value
Hours per week	19.5
Dollars per hour	10.08
Year of dollars	2004
Shift parameter intercept	0.4273
Shift parameter x	0.0183
Shift parameter x^2	-0.0002
Year to begin the shift process	18
Annual probability that someone reattaches to some else following death of spouse	10%

To compute the household production effect, we begin with the following equation:

$$(4.4) \quad H_a = HOURS * \$HOUR * 52 * \Pr SHIFT_a * INFLATION$$

Not all of the value of lost household production will be shifted to others if a person dies or is disabled. 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 *PrSHIFTa*, 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 Bureau of Labor Statistics described above. The results of this estimation and are computed with the following equation:

$$(4.5) \quad \Pr SHIFT_a = \frac{FHH_a}{(HH_a - GQ_a)}$$

The probability of shifting household production *PrSHIFTa* in the event of a disorder is given by the total number of people in households with family members (*FHHa*) divided by the total number of people in households (*HHa*) (less those living in group quarters (*GQa*)). 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 the following equation:

$$(4.6) \quad \$HP_{ty} = \sum_{p}^{P} H_{p+y-1} * (1+ER)^{y-1} * EE_t * PP_{tp} * -1$$

In this equation, *$HPty* is the annual cash flow of shifted household production in year y, where y is the number of years following participation in a program.

4.2 Valuation of Crime Outcomes

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

The current version of WSIPP'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.[47] To model avoided crime costs from these two

[47] 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. Future versions of this model may incorporate some of these additional cost categories.

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.

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 for 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.

4.2a Per-Unit Crime Costs

In WSIPP'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.[48] 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 WSIPP'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.

Exhibits 4.10 and 4.11 display WSIPP's benefit-cost parameters for per-unit costs for the 11 sectors and seven types of crime modeled. The estimates for each row in the Exhibits are described below, along with the sources of the per-unit costs and the uncertainty around the estimates.

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

Exhibit 4.10
Marginal Operating Costs by Crime Type

Resource	Murder	Felony sex crimes	Robbery	Agg-ravated assault	Felony property	Felony drug	Misde-meanor	Year of dollars (of data)	Annual real escalation rate
Police	670	670	670	670	670	670	670	2009	0.027
Courts	152,378	18,770	9,865	4,877	201	201	201	2009	0.020
Juvenile local detention	20,293	20,293	20,293	20,293	20,293	20,293	20,293	2009	0.057
Juvenile local supervision	5,200	5,200	5,200	5,200	5,200	5,200	5,200	2008	0.000
Juvenile state institution	36,743	36,743	36,743	36,743	36,743	36,743	36,743	2009	0.016
Juvenile state supervision	3,927	3,927	3,927	3,927	3,927	3,927	3,927	2009	0.000
Adult jail	21,469	21,469	21,469	21,469	21,469	21,469	21,469	2009	0.022
Adult local supervision	2,877	2,877	2,877	2,877	2,877	2,877	2,877	2013	0.064
Adult state prison	13,422	13,422	13,422	13,422	13,422	13,422	13,422	2014	0.003
Adult post-prison supervision	2,877	2,877	2,877	2,877	2,877	2,877	2,877	2013	0.064
Victim (tangible costs)	567,639	4,745	5,950	12,023	2,027			2010	
Victim (intangible costs)	6,497,488	169,294	8,975	18,567				2010	

Exhibit 4.11
Capital Costs for Crime Resources

Resource	Capital cost per unit	Year of dollars	Finance years
Police	n/a	n/a	n/a
Courts	370	2006	20
Juvenile local detention	200,000	2009	25
Juvenile local supervision	n/a	n/a	n/a
Juvenile state institution	150,000	2009	25
Juvenile state supervision	n/a	n/a	n/a
Adult jail	150,000	2009	25
Adult local supervision	n/a	n/a	n/a
Adult state prison	113,339	2007	25
Adult post-prison supervision	n/a	n/a	n/a
Victim (tangible costs)	n/a	n/a	n/a
Victim (intangible costs)	n/a	n/a	n/a

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. All of these cost parameters are shown in Exhibit 4.11.

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 US 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.

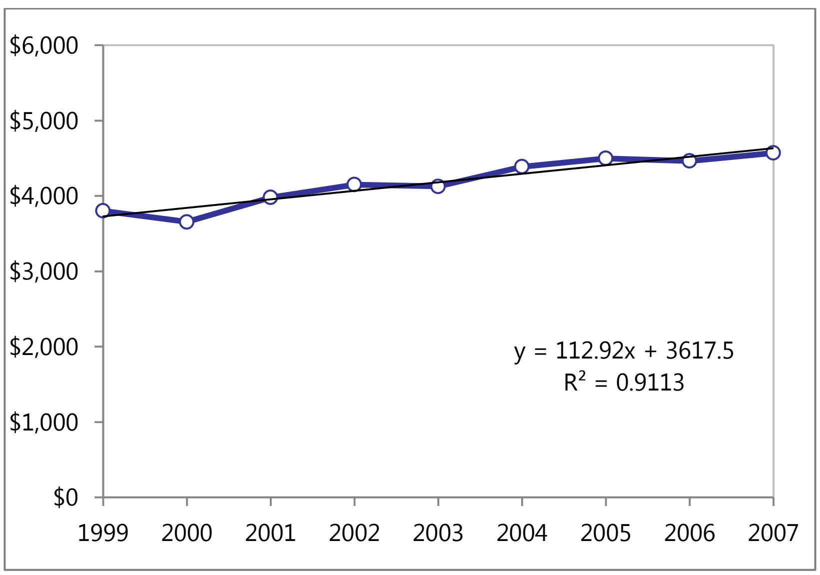
We also collected arrest information for Washington police agencies from the National Archive of Criminal Justice Data maintained by the University of Michigan.[49] Data were collected for calendar years 1994 to 2007, the earliest and latest years available as of December 2009. Arrest data for 1993 were 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 4.12
Average Police Costs per Arrest, 2009 Dollars
Calendar Years 1999 to 2007



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 Exhibit 4.12) 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, as given by the following equation:

$$(4.7)\ Rate = (FV/PV)^{1/N}$$

[49] US 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.

The annual rate of real escalation is 0.027. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.10.

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 seven 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 include 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.[50] 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 0.34). In first-differences, on the other hand, the IPS test indicated a lack of a unit root (IPS p-value 0.000).
- For the two right-hand side variables, the IPS tests indicated a lack of a unit root for A_TOT (IPS p-value of 0.000), but a unit root for TRAFFIC (IPS p-value of 0.88).
- 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 4.13. The sum of the arrest lags is $670. An identical model, but without including a right-hand side dependent variable, produced quite similar results.

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

Exhibit 4.13

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		

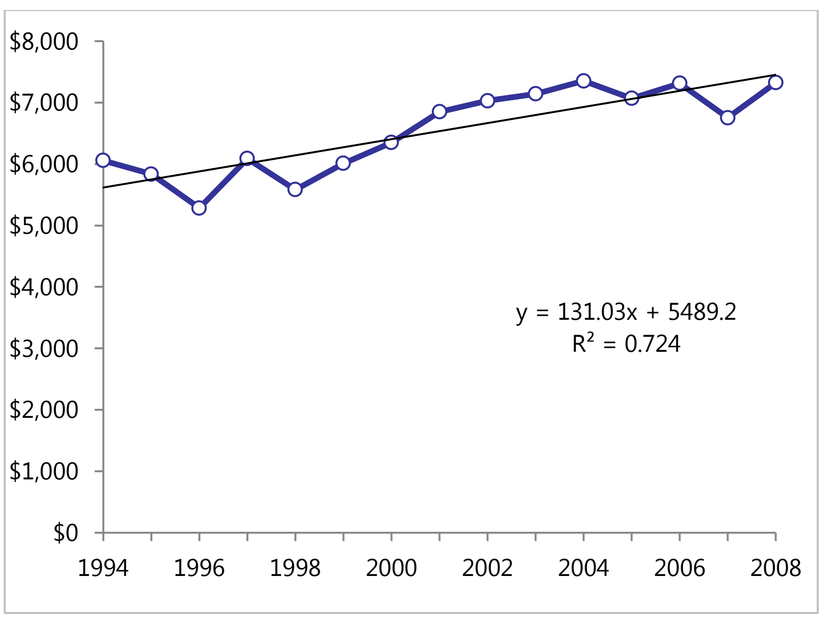
Superior Courts and County Prosecutors Per-Unit Costs. This section describes the steps we use to estimate marginal annual operating costs, and the long-run rate of change in these costs, of county superior courts and prosecutors in Washington State. Our focus is the cost of obtaining convictions in courts, so we combine 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 Exhibits 4.10 and 4.11.

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 includes 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 US 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 4.14
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 Exhibit 4.14) 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 4.7, where FV is the 2008 estimated cost, PV is the 1994 estimate, and N is 14 years.

The annual rate of real escalation is 0.020. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.10.

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 ten 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), and convictions for all non-violent offenses (C_NONVIOL). If unit roots are present, then a simple regression in levels can produce spurious results.[51] 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 0.0028 for M_COURTALLR, 0.0000 for C_HOM, 0.0000 for C_SEX, 0.0000 for C_ROB, 0.0000 for C_ASSLT, 00.0006 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.

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.

[51] Ibid., p. 636.

Exhibit 4.15

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 US Bureau of Justice Statistics annual survey: Justice Expenditure and Employment Extracts.[52] Local government court expenditures in Washington were reported as $19,144,000 for 2006.[53] 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 4.11, 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 4.8, 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, as given by the following equation:

$$(4.8) \quad PMT = \frac{iPV}{1-(1+i)^{-n}}$$

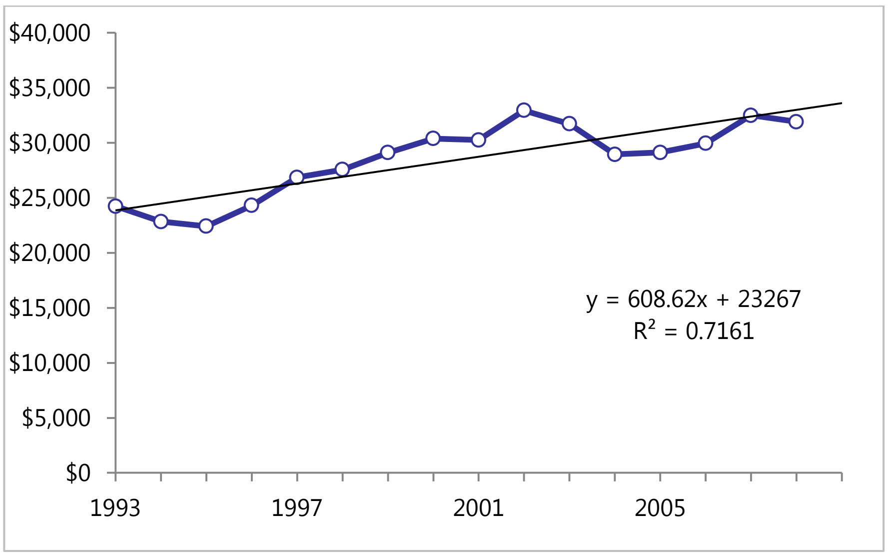
Local Adult Jail Per-Unit Costs. This section describes the steps we use to estimate marginal annual jail operating costs, and the long-run rate of change in these costs, of the county-run adult jail system in Washington State. We also describe our estimate of the capital cost per jail bed. All of these cost parameters are entered into the crime model, as shown in Exhibit 4.10. In WSIPP'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. WSIPP assumes the same annualized per-day jail cost for both these events.

[52] 2006, published December 1, 2008 (NCJ 224394).
[53] Ibid. Table 4, Justice system expenditure by character, State and type of government, fiscal 2006.

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, electronically 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 US 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 jail average daily 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 4.16
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 in Exhibit 4.16) 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 4.7, where FV is the 2008 estimated cost, PV is the 1993 estimate, and N is 15 years.

The annual rate of escalation is 0.022. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.10.

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.[54] 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 0.713). In first-differences, the test indicated a lack of a unit root (IPS p-value 0.000).

[54] Ibid., p. 636.

- 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 0.975). With time trends included, the IPS test continued to indicate a unit root (p-value 0.582). In first-differences, the test indicated a lack of a unit root (IPS p-value 0.000).
- 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.

Since the two series have unit roots, we tested to determine if the two series together are cointegrated.[55] We used two versions of a panel cointegration test in EVIEWS. Both the Pedroni Engle-Granger test (p-value 0.000) and the Kao Engle-Granger test (p-value 0.000) 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 0.113. 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.[56]

[55] Ibid., p. 639.
[56] 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.

Exhibit 4.17

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
JAILRREAL(-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 in 2009 dollars per county jail bed. In our crime model, the total capital cost per bed is converted to an annualized capital payment, with equation 4.8, 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.

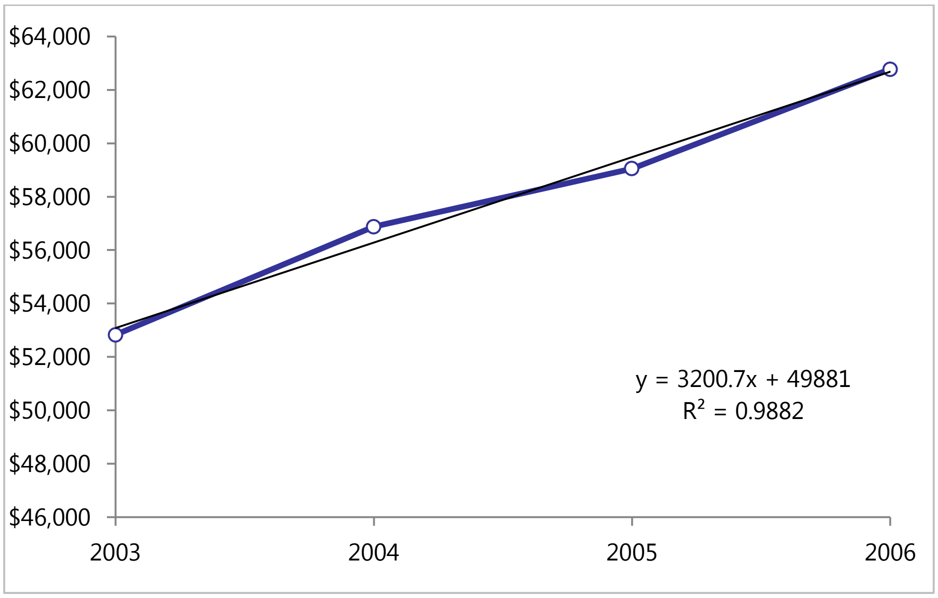
Local Juvenile Detention and Probation Per-Unit Costs. This section describes the steps we use to estimate marginal annual detention operating costs, and the long-run rate of real (inflation-adjusted) change in these costs of county-run juvenile detention facilities in Washington. We also describe our 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 4.10.

Detention Operating Costs. For an estimate of the marginal operating cost of state juvenile offender institutions, we conduct 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 electronically available, as of winter 2010. The Auditor's data for the expenses include the categories for residential care and custody (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 US 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 ADP of the facilities statewide.

We computed the average costs per institutional ADP (in 2009 dollars) and plotted these data in Exhibit 4.18.

Exhibit 4.18
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 Exhibit 4.18) 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 4.5, where FV is the 2006 estimated cost, PV is the 2003 estimate, and N is three years.

The annual rate of real escalation is 0.057. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.10. 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 4.19

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 in 2009 dollars.[57] In our crime model, the total capital cost per arrest is converted to an annualized capital payment, with equation 4.8, 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. As far as we know, there is not a data source for the average daily juvenile court probation caseloads in Washington.

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.

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

- 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 a WSIPP 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.[58]
- We then estimated the 2008 average daily probation caseload of juvenile courts as 3,207 (5,660, multiplied by 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 from DOC, the ratio of marginal costs to average costs was 0.50 (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 0.50 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 4.10.

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 4.10.

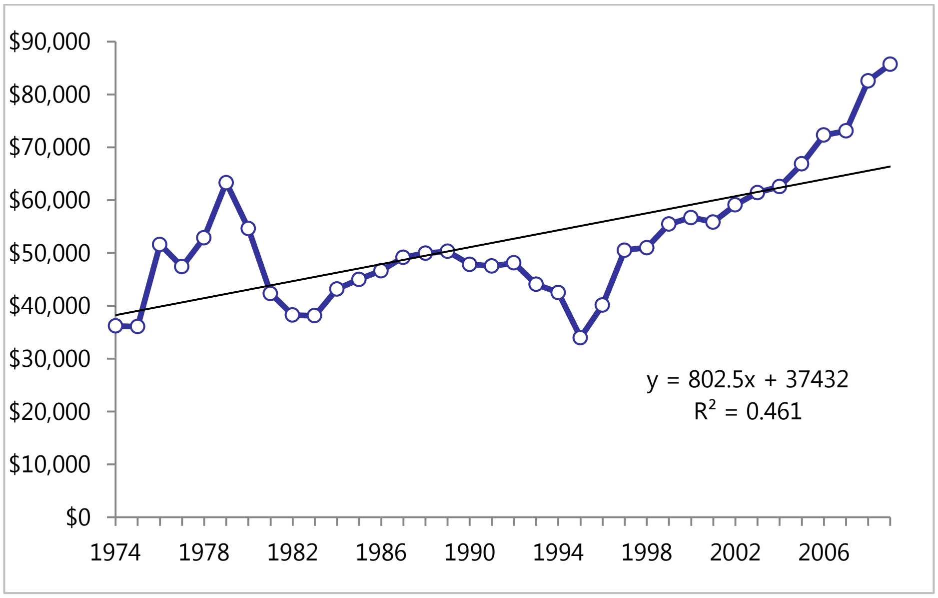
Institutional Operating Costs. For an estimate of the marginal operating costs of state juvenile offender institutions, we conducted a time-series analysis of annual data for institutional expenditures and average daily institutional population for JRA for fiscal years 1974 to 2009. The expenditure data 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 US 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 4.20.

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

Exhibit 4.20

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 Exhibit 4.20) 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 4.5, where FV is the 2009 estimated cost, PV is the 1974 estimate, and N is 35 years.

The annual rate of escalation is 0.016. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.10. The data plotted in Exhibit 4.20 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 0.016 may be on the low side if 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.[59] 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 0.511 without a time trend and 0.620 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 0.299 without a time trend and 0.760 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.[60] 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 0.01) for the null hypothesis that the residual series has a unit root. Therefore, this

[59] Wooldridge, (2009), p. 636.
[60] Ibid., p. 639.

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.

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

Exhibit 4.21

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.[61] The three ADP variables were jointly significant with a p-value on the F test of 0.0473. 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 2009 dollars. In our crime model, the total capital cost per arrest is converted to an annualized capital payment, with equation 4.8, 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

[61] 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.

estimated the marginal expenditure per average annual caseload. We found, from our time-series analysis of the community supervision costs of DOC, the ratio of marginal costs to average costs was 0.50. Multiplying $7,847 by 0.50 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 4.10.

State Department of Corrections (DOC) Per-Unit Costs. This section describes our estimates for the Washington DOC's 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. Unlike other DOC cost estimates, the marginal cost of a prison bed is a negotiated price. DOC's budget staff calculates a marginal cost estimate prior to each legislative session. A meeting is held with DOC budget staff, legislative fiscal analysts from the Senate Ways and Means and the House Appropriations Committees, a fiscal analyst from the Office of Financial Management, and WSIPP staff, to negotiate the marginal cost that will be used for the legislative session. Exhibit 4.22 displays the marginal cost estimates for each legislative session. Thus, our benefit-cost model currently uses the marginal estimate of $13,422.

Exhibit 4.22

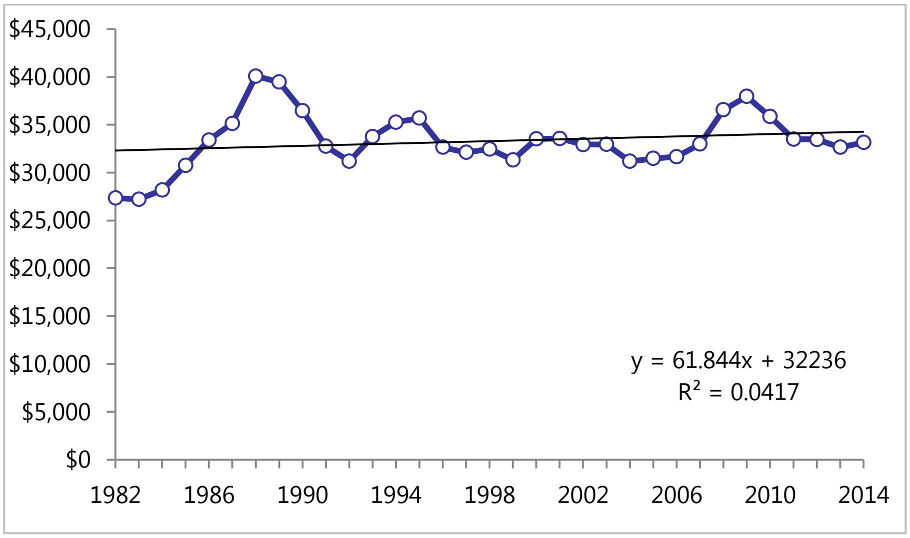
DOC Average Daily Prison Bed Marginal Cost Estimate,
2014 Dollars

Legislative session	Marginal cost per prison bed
2017	$13,422
2016	$13,563
2015	$12,216
2014	$11,966
2013	$11,536

For comparison purposes, we analyzed annual data for DOC institutional expenditures and average daily prison population for fiscal years 1982 to 2014. The expenditure data were obtained from 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 US 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 2014, 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 2014 dollars) for 1982 to 2014 and plotted the results below. The annual rate of real escalation in average costs is 0.003. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.10.

Exhibit 4.23
Average DOC ADP Prison Costs, 2014 Dollars
Fiscal Years 1982 to 2014



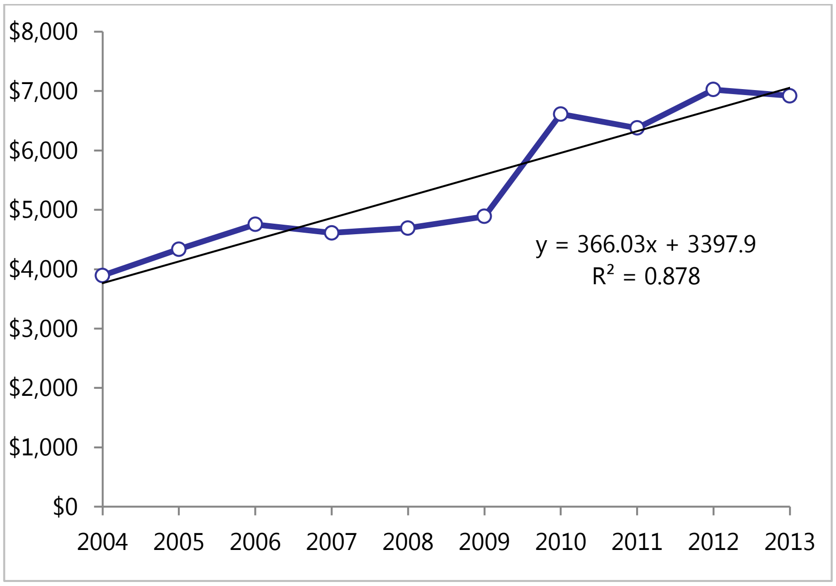
Prison Capital Costs. DOC capital costs for new institutional beds were estimated. Capital cost estimates for the relatively new Coyote Ridge medium security facility in Washington were obtained from legislative fiscal staff. The 2,048 bed facility cost $232,118,000 (thus, a per-bed cost of $113,339) 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 4.11. In our crime model, the total construction costs per-bed are converted to an annualized capital payment, with equation 4.8, 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 DOC's 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 2004 to 2013. The expenditure data were obtained from LEAP for Agency 310 (Department of Corrections) for code 300 (community supervision).[62] 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 2013 dollars using the aforementioned price index. The average community supervision cost over the 2004 to 2013 period is $4,927.

[62] Expenditure data are available back to 1998; however, due to an accounting change, we use data from 2004 through 2013 to estimate costs using a consistent data series.

Exhibit 4.24

Average DOC Average Daily Community Supervision Costs,
2013 Dollars, Fiscal Years 2004 to 2013



Over the 2004 to 2013 period, there was a significant upward trend in the inflation-adjusted per-unit costs, as revealed by the linear regression line shown in Exhibit 4.24. 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 2004 ($3,764) and 2013 ($6,692) and calculated the annual rate of escalation for the nine years using equation 4.7 where FV is the cost estimate for 2013, PV is the estimate for 2004, and N is nine 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 4.10. 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.[63] 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 0.366 without a time trend, and was not significant at 0.655 with a time trend. In first-differences, the ADF p-value for the DOCCSREAL series was 0.3622, indicating a unit root in a first-differenced data series.
- For the DOCCSADP series, the p-values for the ADF test were 0.8117 without a time trend and 0.6320 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 0.0997 indicating 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.

[63] Wooldridge, (2009), p. 636.

Assuming the two series have unit roots, we used an Engle-Granger test to determine if the two series together are cointegrated.[64] The test failed to reject the null hypothesis with a p-value of 0.5577, meaning the series are not cointegrated. Thus, we proceed with a first-difference model as our preferred approach. We also include a lagged first difference dependent variable in our regression model because we expect that prior spending influences current changes in spending.

The sum of the two DOCCSADP coefficients total $2,877 per ADP, in 2013 dollars. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.10.

Exhibit 4.25

Dependent Variable: DOCCSREAL2-DOCCSREAL2(-1)
Method: Least Squares
Date: 05/04/15 Time: 17:07
Sample (adjusted): 2006 2013
Included observations: 8 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3646346.	1446850.	2.520197	0.0653
DOCCSREAL2(-1)-DOCCSREAL2(-2)	0.437458	0.185275	2.361120	0.0776
DOCCSADP2-DOCCSADP2(-1)	1725.168	361.9286	4.766598	0.0089
DOCCSADP2(-1)-DOCCSADP2(-2)	1151.991	417.7637	2.757518	0.0510

R-squared	0.915265	Mean dependent var	-2519899.
Adjusted R-squared	0.851714	S.D. dependent var	7673875.
S.E. of regression	2955052.	Akaike info criterion	32.94278
Sum squared resid	3.49E+13	Schwarz criterion	32.98250
Log likelihood	-127.7711	Hannan-Quinn criter.	32.67488
F-statistic	14.40200	Durbin-Watson stat	2.090707
Prob(F-statistic)	0.013076		

This first-difference model is our preferred model. Our model in levels revealed a negative relationship between the prior period 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, while 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. After a review of the literature, we chose to use the average of victim cost estimates from two papers, McCollister (2010) and Cohen & Piquero (2009), in WSIPP's benefit-cost model with some modifications.[65] These crime victim costs build on and modify the previous work prepared for the US Department of Justice by Miller, Cohen, and Wiersema (1996).[66]

[64] Ibid., p. 639.
[65] 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. Cohen, M.A., & Piquero, A.R. (2009). New evidence on the monetary value of saving a high-risk youth. *Journal of Quantitative Criminology*, *25*(1), 25-49.
[66] 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.

The McCollister study divides 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 WSIPP's model, we only include McCollister's tangible victim costs because 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 lacking. This FBI reporting practice requires the adjustment made by McCollister. For application to WSIPP's benefit-cost model, however, this adjustment is not necessary. WSIPP'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 4.2c of this Chapter for a description of WSIPP's data sources for counting convictions.

The Cohen study reports one number for victim costs of crime for each type of crime. WSIPP combines the two types of robbery reported in the Cohen paper to better match the crime types used in the model. We apply the percentage breakout of tangible and intangible costs from the McCollister paper to the average of total victim costs for the two papers.

WSIPP's model also has one crime category for felony property crimes. Both the McCollister and Cohen study break 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.

WSIPP's modified crime victim cost estimates are included in the crime model, as shown in Exhibit 4.10.

Sources of per-Unit Costs and Uncertainty around Estimates. Exhibit 4.26 shows the breakouts and sources of criminal justice costs for Washington State. The model also calculates with an estimated uncertainty around each resource-specific unit cost. This uncertainty is used when Monte-Carlo simulations are run in the model. For most criminal justice costs, the model uses a per unit cost variance of plus or minus 0.1. The variation in WSIPP crime victim cost estimates is calculated as the variation of total victim crime costs for each crime type between the two studies weighted by the number of crimes of each crime type for Washington and is equal to 0.08.

Exhibit 4.26

Proportional of Marginal Criminal Justice Costs by Funding Source

	Operating			Capital		
	State	Local	Federal	State	Local	Federal
Police[1]	15%	85%	0%	22%	78%	0%
Courts & prosecutors[2]	16%	84%	0%	21%	79%	0%
Juvenile local detention	20%[3]	80%	0%	0%[4]	100%	0%
Juvenile local supervision	13%[3]	87%	0%	0%[4]	100%	0%
Juvenile state institution[5]	100%	0%	0%	100%	0%	0%
Juvenile state supervision[5]	100%	0%	0%	100%	0%	0%
Adult jail	11%[5]	89%	0%	0%[4]	100%	0%
Adult local supervision	27%[5]	73%	0%	0%[4]	100%	0%
Adult state prison[6]	100%	0%	0%	100%	0%	0%
Adult post prison supervision[6]	100%	0%	0%	100%	0%	0%

[1] Justice Expenditure and Employment Extracts, 2010—Preliminary, Tracey Kyckelhahn, Ph.D., Tara Martin, BJS Intern, July 1, 2013. NCJ 242544, Table 4: Justice system expenditure by character, state and type of government, fiscal 2010, available at: http://www.bjs.gov/index.cfm?ty=pbdetail&iid=4679. Direct current Police Protection expenditures for state and local governments for Washington State.
[2] Justice Expenditure and Employment Extracts, 2010—Preliminary, Tracey Kyckelhahn, Ph.D., Tara Martin, BJS Intern, July 1, 2013. NCJ 242544, Table 4: Justice system expenditure by character, state and type of government, fiscal 2010, available at: http://www.bjs.gov/index.cfm?ty=pbdetail&iid=4679. Direct current Judicial and Legal expenditures for state and local governments for Washington State.
[3] Sources for operating costs of juvenile local detention and juvenile local supervision estimated by analyzing information from the Washington State Auditor's Local Government Finance Reporting System (LGFRS) system. (Functional Group/BARS Summary, Expenditures for government types City/Town and County, All Objects, All Available Fund Types, for 2011). http://portal.sao.wa.gov/LGCS/Reports/, Juvenile Services (BARS account: 527).
[4] WSIPP assumes capital costs for all local juvenile and adult resources are 100% locally funded.
[5] Sources for operating costs of adult jail and adult local supervision estimated by analyzing information from the Washington State Auditor's Local Government Finance Reporting System (LGFRS) system. (Functional Group/BARS Summary, Expenditures for government types City/Town and County, All Objects, All Available Fund Types, For 2011). http://portal.sao.wa.gov/LGCS/Reports/, Detention and Correction (BARS account: 523).
[6] WSIPP assumes all state funded.

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 4.27 displays how criminal justice resources in Washington State are used in response to crime. The estimates for each row in the Exhibit are described below.

Probability of Resource use. The first block of information in Exhibit 4.27 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% chance the offender will receive a prison sentence and a 29% chance of receiving a jail sentence. These sentencing probabilities were obtained from the Washington State Sentencing Guidelines Commission.[67]

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.

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

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 WSIPP.[68]

Juvenile Local Supervision. The average length of stay on probation was also estimated from the same survey of juvenile courts conducted by WSIPP.[69]

Juvenile State Institution. The average length of stay in a juvenile state institution was estimated using data obtained from the Sentencing Guidelines Commission.[70]

Juvenile State Supervision. The average length of stay on juvenile parole was estimated using information obtained from the Juvenile Rehabilitation Administration.[71]

Adult Jail, With Local Sentence. The average length of stay in jail for local sentences was estimated using data from the Sentencing Guidelines Commission.[72]

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.[73]

Adult Community Supervision and Adult Post-Prison Supervision. These numbers were obtained from the Sentencing Guidelines Commission.[74]

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 4.27 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.

[68] Burley & Barnoski, (1997).
[69] Ibid.
[70] Washington State Sentencing Guidelines Commission (personal communication, March 10, 2010).
[71] Washington State Juvenile Rehabilitation Administration (personal communication, April 18, 1997).
[72] Sentencing Guidelines Commission, (2009), Table 1.
[73] Washington State Department of Corrections (personal communication, November 7, 1996).
[74] Washington State Sentencing Guidelines Commission (personal communication, April 6, 2010).

Exhibit 4.27

Use of Crime Resources by Crime Type

Resource	Murder	Felony sex crimes	Robbery	Aggravated assault	Felony property	Felony drug	Misdemeanor
Probability of resource use, given a crime (by type of crime)							
Police	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Courts	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Juvenile local detention	0.14	0.54	0.32	0.66	0.85	0.86	0.98
Juvenile local supervision	1.00	1.00	1.00	1.00	1.00	1.00	0.00
Juvenile state institution	0.86	0.46	0.68	0.34	0.15	0.14	0.02
Juvenile state supervision	1.00	1.00	1.00	1.00	1.00	1.00	0.00
Adult jail	0.01	0.31	0.29	0.58	0.63	0.67	1.00
Adult local supervision	1.00	0.85	0.89	0.69	0.17	0.73	0.00
Technical violation -local supervision	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Adult state prison	1.00	0.70	0.71	0.42	0.37	0.33	0.00
Adult post-prison supervision	0.99	0.96	0.98	0.81	0.29	0.96	0.00
Police	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Courts	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Juvenile local detention, for local sentence	0.04	0.04	0.04	0.04	0.04	0.04	0.00
Juvenile local detention, for state sentence	0.02	0.02	0.02	0.02	0.02	0.02	0.00
Juvenile local supervision	0.57	0.57	0.57	0.57	0.57	0.57	0.57
Juvenile state institution	1.65	0.90	0.96	0.67	0.53	0.63	0.19
Juvenile state supervision	0.47	1.49	0.44	0.45	0.48	0.55	0.47
Adult jail, for local sentence	1.00	0.58	0.50	0.34	0.24	0.23	0.10
Adult jail, for prison sentence	1.08	0.48	0.44	0.37	0.32	0.28	0.00
Adult local supervision, jail sentence	2.00	2.50	1.01	0.82	0.24	0.90	0.50
Technical violation -local supervision	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Adult state prison	16.43	5.82	4.57	2.65	1.65	1.36	0.00
Adult post-prison supervision	3.91	3.70	2.94	1.67	0.51	1.06	0.00
Adult	0.18390	0.18390	0.18390	0.18390	0.18390	0.18390	0.18390
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	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. The amount of time actually served is often shorter than the original sentence. Exhibit 4.27 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.

To account for these lengthening sentences, we use the sentencing grid and WSIPP'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.

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 are released from prison have at least three prior offenses and then determine the following:

- likelihood of conviction,
- likelihood of going to prison if convicted, and
- 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.

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 portion of a year. As of the date of this publication, 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 4.27 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 4.27 as representative of the typical decisions made pursuant to current Washington State law.

4.2b Criminological Information for Different Populations

To estimate the long-run impacts of evidence-based programs on crime, WSIPP 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 (50 years).

Crime Parameters. Exhibit 4.28 presents an example of the calculations we perform to determine 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 4.2 of this Chapter.

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 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 4.28
Crime Parameters from Example Population: Adult Supervision (General)

Population	Number of follow up years	Cumulative recidivism/crime over the period			Hazard rate: timing of any recidivism/crime		
Adult supervision—General	15	Exponential distribution 0.24376 0.14270 -0.02078 0.00141 -0.00004			4th order polynomial 0.15844 -0.03186 0.00410 -0.00026 0.00001		
Crime base population parameters	**Murder**	**Felony sex offenses**	**Robbery**	**Aggravated assault**	**Felony property**	**Felony drug**	**Misde-meanor**
Percent where most serious recidivism or crime offense is:	0.011	0.024	0.046	0.185	0.23	0.16	0.344
Number of trips, where most serious is:	1.015	1.067	1.421	1.516	3.01	4.24	7.891
Average number of offenses per trip	1.256	1.490	1.337	1.268	1.40	1.23	1.133

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.[75] 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 WSIPP's criminal history database, which is a synthesis of information for offenders convicted in Washington State.[76]

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.[77] 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.

---

[75] Barnoski, R. (1997). *Standards for improving research effectiveness in adult and juvenile justice. (*Doc. No. 97-12-1201). Olympia: Washington State Institute for Public Policy, p. 2.
[76] Criminal history data are from the Washington State Administrative Office of the Courts and Department of Corrections.
[77] See Barnoski R., & Drake, E. (2007). *Washington's offender accountability act: Department of Corrections' static risk instrument* [Revised October, 2008] (Doc. No. 07-03-1201). Olympia: Washington State Institute for Public Policy. See also, Barnoski, R. (2004). *Assessing risk for re-offense: Validating the Washington State juvenile court assessment.* (Doc. No. 04-03-1201). Olympia: Washington State Institute for Public Policy.

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 WSIPP'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 for the full 50-year period.

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.[78] 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 4.29). Poverty is measured as less than 200% of the federal poverty threshold. The coefficient from this model indicates that poverty is significantly related with a greater likelihood of crime ($b = 0.803$, $p < 0.0001$). 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,

$$= (\exp(0.803) * 0.33)/(1 - 0.33 + 0.33 * \exp(0.803)).$$

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 take the ratio of the odds ratios for each of those crime categories relative to the total poverty effect and multiply the ratio of odds ratios by the crime probability for the non-offender population and normalize the crime probability to one.

Exhibit 4.29

Effect of Poverty on Arrests

	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 $p < 0.001$.

[78] US 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.

4.2c Estimates of Victimizations per Conviction

Nearly all of the effect sizes computed from programs and policies impacting crime describe official measures of criminal activity, such as convictions or arrests. There may be, of course, many more criminal victimizations than those reported in official measures of crime. To estimate the number of victimizations per officially reported crime, WSIPP's benefit-cost model uses additional information. Parameters displayed in Exhibit 4.30 are described below.

Exhibit 4.30
Estimated Victims per Conviction

FBI UCR data	Murder	Rape	Robbery	Aggravated assault	Burglary	Theft	Motor vehicle theft	Year of data
Number of statewide crimes reported to police	159	2,238	5,579	11,766	55,597	160,446	24,384	2011
Multiplicative adjustment to align with felonies	1.000	2.410	1.000	1.000	1.000	0.235	1.000	
Victimization numbers								
Calculated adjusted crimes reported to police	159	5,394	5,579	11,766	55,597	37,705	24,384	
Percent of crime reported to police	1.0	0.307	0.660	0.670	0.520	0.685	0.830	2011
Calculated estimate of statewide felony crimes	159	17,569	8,453	17,561	106,917	55,044	29,378	
Washington court and criminal justice numbers	**Murder**	**Felony sex crimes**	**Robbery**	**Aggravated assault**	**Felony property**			
Number of arrests, adult and juvenile	148	1,918	1,892	5,456	35,819			
Statewide number of convictions, adult and juvenile	204	1,604	908	4,299	9,005			
Number of counts, adult and juvenile	286	3,273	1,432	7,054	18,937			
Average number of offenders per victim	1.00	1.00	1.00	1.00	1.00			
Percent of other crimes per conviction	0.64	0.20	0.20	0.20	0.20			
Estimated victimizations per convicted offender	**1.00**	**3.82**	**3.12**	**2.13**	**5.93**			

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 victims[79] and other sexual assaults.[80] 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.[81]

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.[82]

---

[79] Bureau of Justice Statistics. (2008). C*riminal 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.
[80] Ibid., Table 1.
[81] Snyder, H.N. (2000). *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.
[82] Bureau of Justice Statistics, (2008), Table 100.

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.[83] We adjust the percentage of crimes reported to police from the NCVS for sex offenses and theft offenses to reflect the multiplicative adjustment to align UCR data with Washington felonies.

Statewide Number of Convictions, Adult and Juvenile. Adult and juvenile felony conviction data are obtained from the Administrative Office of the Courts.[84]

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).[85] 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. 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%. 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.

Variance in Ratios of Victimizations per Convicted Offender. Since the previous parameter, the percentage of other crimes per conviction, is estimated with considerable imprecision, in Monte Carlo simulation the parameter is selected from a triangular probability distribution that bounds risk. The user can enter a low and high percentage that is then applied to the parameter. We have chosen 20% lower and 20% higher bounds for the triangular distribution.

4.2d Procedures to Estimate Criminal Justice System and Victimization Events

In this section of the Benefit-Cost Technical Documentation, 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 Exhibits 4.10 and 4.11, we estimate costs avoided using the following equation:

$$(4.9)\quad CjsResource\$_{ry} = \sum_{c=1}^{C} \sum_{t=1}^{ceiling(T_c)} \sum_{f=1}^{F} \left[CjsEvent_{yctf} \times CrimePr_c \times CjsResourcePrW_{rc} \times TripPr_{ct} \times TimetoRecid_f \times Unit\Delta_y \times CjsResourceCost_{rc} \right] \times RecidRate$$

We also count Average Daily Population prison beds avoided. We do this using equation 4.9 above however; we do not multiply by the $CjsResourceCost_{rc}$.

Variable Definitions. Below are definitions and calculations for the variables used in equation 4.9.

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.

---

[83] Ibid.

[84] 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

[85] US 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.

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 0.08. See also $TripPr_{ct}$.

$CrimeEvent_{yctf}$—Variable indicating if and when a criminal justice resource is used (shown in Exhibit 4.27) or whether a victimization occurs and, if so, how much of the criminal justice system resource is used. For each criminal justice system resource or victimization, we calculate an event matrix, $CrimeEvent_{yctf}$, to indicate when a resource is used. Each event matrix occurs within the recidivism follow-up period, *f*, for each trip, *t*, and within the 50-year follow-up, *y*. We estimate this matrix for each crime type, *c*. For criminal justice system events that occur over multiple years (e.g., prison), we incorporate length of stay information from Exhibit 4.27into the event matrix.

$CrimePr_c$ —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 4.28.

$CjsResourcePrW_{rc}$.—The probability that a criminal justice resource, *r*, will be used for a specific type of crime, c. See Exhibit 4.27. 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 one.

$TripPr_{ct}$—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 4.28. The probability of a trip occurring is one. Once a whole trip has been used, then we use the remaining probability of the trip. For example, adult offenders under general supervision whose most serious reoffense is a sex offense have an average number of 1.07 trips in a 15-year follow-up period. Thus, there is a probability of one trip occurring and a probability of 0.07 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.

$TimetoRecid_f$—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.

$Unit\Delta_y$—The change in the probability of being an offender (vs. not being an offender) in year y.

$CjsResourceCost_{rc}$—The per unit marginal costs of each criminal justice resource as estimated in Section 4.2a of this Chapter and as shown in Exhibits 4.10 and 4.11.

$RecidRate$—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 4.28. Different recidivism base rates are used depending on the specific population that receives a given program. See Exhibit 4.28.

Victimizations Avoided. Using information from Exhibit 4.28 and 4.30, we estimate the number of victimizations avoided and victimization costs avoided using the following equation:

$$(4.10)\quad Victim\$_{ry} = \sum_{c=1}^{C} \sum_{t=1}^{ceiling(Trips_c)} \sum_{f=1}^{F} \left[CrimeEvent_{yctf} \times VictimVolume_{yctf} \times CrimePr_c \times TripPr_{ct} \times_f Unit\Delta_y \times VictimCost_{rc}\right] \times RecidRate$$

Variable Definitions. Below are definitions and calculations for the variables used in equation 4.10 unless otherwise defined in the aforementioned section.

$VictimVolume_{ctf}$—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 two.

$$(4.11)\quad VictimVolume_{cf} = \sum_{c=1}^{C}\sum_{v=c}^{C}\sum_{f=1}^{F}\frac{(MostSeriousTripVic_c + AddVicsMostSeriousTrip_c + RemainingTrips_c)}{Trips_c}$$

Equations 4.11, 4.12, and 4.12 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.

$$(4.12)\quad MostSeriousTripVic_c = 1 \times VicsPerConvictedOffender$$

$$(4.13)\quad AddVicsMostSeriousTrip_c = OffensesPerTrip_c \times VicsPerConvictedOffender_v \times \left(\frac{CrimePr_v}{\sum_{c=1}^{C} CrimePr_c}\right)$$

$$(4.14)\quad RemainingTrips_c = (Trips_c - 1) \times OffensesPerTrip_c \times VicsPerConvictedOffender_c \times \left(\frac{CrimePr_v}{\sum_{c=1}^{C} CrimePr_c}\right)$$

$CrimeEvent_{yctf}$—A dichotomous variable indicating if a victimization event has occurred during the at-risk year. Victimizations are shown in Exhibit 4.30.

$VictimCostrc$—The per-unit cost of crime to victims as estimated in Section 4.2 of this Chapter and as shown in Exhibits 4.10 and 4.11.

4.2e Linkages: Crime and Other Outcomes

WSIPP's benefit-cost model monetizes improvements in crime, in part, with linkages between crime 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 juvenile crime and high school graduation 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 parameters are entered into the benefit-cost model and used when performing a Monte Carlo simulation. The linkages in the current WSIPP model are listed in the Appendix.

4.2f Special Calculations for Prison and Policing Resources

How prison incarceration rates affect crime and how the number of police officers affects crime are most often summarized with an "elasticity" effect size metric, rather than a D-cox or Cohen's d effect size metric. This section of the Technical Documentation describes the particular methods we use to estimate effects and monetize outcomes for these two elasticity-based topics.

We conducted a meta-analytic review of the research literature from the US and beyond to determine if prison and police are effective at reducing crime rates. We examine studies that have measured how prison average daily population (ADP) or the number of police officers (POL) affect current crime rates. A fuller explanation of WSIPP's meta-analysis for these two topics are described in a separate WSIPP report.[86]

[86] Aos, S., & Drake, E. (2013). *Prison, police, and programs: Evidence-based options that reduce crime and save money* (Doc. No. 13-11-1907). Olympia: Washington State Institute for Public Policy, available at: http://www.wsipp.wa.gov/rptfiles/13-11-1901.pdf

There is a research literature on the effect of incarceration rates on crime.[87] Many of the studies addressing this relationship in the US construct models using state-level data over a number of years to estimate the parameters of an equation of this general form:

$$(4.15)\quad C_{tsy} = a + b(ADP_{sy}) + c(X_{sy}) + e$$

In this typical model, crime, C, of type, t, 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, often including state and year fixed effects, and an error term, e. Some studies use this type of model to estimate total reported crime, while others examine types of crime such as violent crime or property crime.

There is similar research literature on the effect of the number of police officers on crime rates.[88] Many of these studies use data at the city or county level to estimate the parameters of an equation, such as the following:

$$(4.16)\quad C_{tcy} = a + b(POL_{cy}) + c(X_{cy}) + e$$

In a typical police model, crime, C, of type, t, in city or county, c, and year, y, is estimated to be a function of the size of a city's or county's overall commissioned police force, POL, a vector of control variables, X, often including city/county and year fixed effects, and an error term, e.

In the research literature we reviewed, these models are almost always estimated with a log-log functional form, at least for the dependent and policy variables. Several authors have observed that the panel time series often used to estimate equations 4.15 and 4.16 are likely have unit roots, especially with state level data.[89] Thus, to help avoid estimating spurious relationships, some authors estimate equations 4.15 and 4.16 in first-differences since the time series typically do not exhibit unit roots after differencing once.

There is considerable concern in the research literature on the econometric implications of possible simultaneous relationships between the variables of interest in equations 4.15 and 4.16 and in omitted variables bias.[90] Simultaneity can occur because crime may be a function of ADP or POL, but ADP and POL may also be a function of crime. Failure to account for these simultaneous relationships, as well as failure to address omitted control variables in regressions, can cause statistically biased estimates. In recent years, much of the discussion and debate in the research literature has focused ways to address statistical bias from simultaneity and omitted control variables. In our meta-analyses, we only included studies that met rigorous standards of evidence by accounting for simultaneity.

Meta-Analytic Results. Exhibit 4.31 displays the results of our meta-analyses. The results are shown for both prison and police policy variables and their estimated effects on violent crime and property crime. Exhibit 4.33 displays the meta-analytic results for prison length of stay on criminal recidivism.

[87] Marvell, T.B. (2010). Prison population and crime. In B.L. Benson, & P.R. Zimmerman (Eds.). *Handbook on the Economics of Crime* (pp. 145-183). Cheltenham, UK: Edward Elgar Publishing.

[88] Lim, H., Lee, H., & Cuvelier, S.J. (2010). The impact of police levels on crime rates: A systematic analysis of methods and statistics in existing studies. *Asia Pacific Journal of Police & Criminal Justice, 8*(1), 49-82.

[89] See, for example, Marvell, (2010). See also, Spelman, W. (2008). Specifying the relationship between crime and prisons. *Journal of Quantitative Criminology, 24*, 149-178.

[90] Durlauf, S.N., & Nagin, D.S. (2010). The deterrent effect of imprisonment NBER 5/07/10, downloaded from: www.nber.org/chapters/c12078

Exhibit 4.31

Meta-Analytic Results: Prison ADP and Police Levels on Current Crime Levels

Policy topic & outcome				
Topic	Dependent variable: Type of crime	Elasticity	Standard error	Number of studies
Prison: average daily population	Total	-0.260	0.026	7
	Violent	-0.351	0.095	6
	Property	-0.246	0.029	6
Police: number of officers	Total	-0.377	0.086	9
	Violent	-0.763	0.116	7
	Property	-0.351	0.123	7

Notes: All results are from random effects meta-analyses estimated with the methods described in Chapter 2.

In order to compute benefit-cost estimates, the meta-analyzed elasticities reported on prison and police as reported in Exhibit 4.31 need to be converted into the number of crimes avoided or incurred with a particular change in prison or policing levels.

To begin, the usual calculation of marginal effects from the elasticities obtained with log-log crime models is obtained for the effect of prison on crime (equation 4.17) and for the effect of police on crime (equation 4.18)using the following equations:

$$(4.17)\ \ \Delta C_t = \frac{E_t \times \left(\frac{C_t}{ADP}\right)}{RR_t} \qquad (4.18)\ \ \Delta C_t = \frac{E_t \times \left(\frac{C_t}{POL}\right)}{RR_t}$$

In equations 4.17 and 4.18, the change in the number of crimes, ΔC, for a particular type of crime, t, is estimated with: 1) E, the crime-prison elasticity or the crime-police elasticity for a particular type of crime, t, obtained from the relevant meta-analysis reported in Exhibit 4.31; 2) the reported level of crime, C, for a particular crime type, t, as reported in Exhibit 4.30; 3) the incarceration rate, ADP, or the level of police employment, POL, as reported in Exhibit 4.27; and 4) the reporting rate to police by crime victims, RR, for a particular type of crime, t, as reported in Exhibit 4.30. In many studies, the marginal effects are often calculated at the mean values for ADP, POL, C_t, and RR_t over the time series. For policy purposes, however, it is more relevant to use more recent values for these variables.

As noted earlier, the UCR definition of Part 1 crimes may not match a state's current definition of felony crimes. Therefore, we make adjustments to the reported UCR crimes for two types of crimes, sex offenses and larceny/theft (see our adjusted inputs in Exhibit 4.30), to more closely align the UCR definitions with current law definitions in Washington, using the following equation:

$$(4.19)\ \ C_t = UCR_t \times UCRAdj_t$$

In this analysis, we implement equations 4.17 and 4.18 for two types of crime: violent crime and property crime. Additionally, to address the limitations in the policy-relevance of the overall elasticities, we implement two adjustments to the meta-analyzed elasticities, E_t, on prison and police as reported in Exhibit 4.31. Therefore, we modify equations 4.17 and 4.18 as follows:

$$(4.20)\ \ \Delta C_v = \frac{(E_v \times R_v \times P_v) \times \left(\frac{C_v}{ADP}\right)}{RR_v} \qquad (4.21)\ \ \Delta C_v = \frac{(E_v \times R_v \times P_v) \times \left(\frac{C_v}{POL}\right)}{RR_v}$$

$$(4.22)\ \ \Delta C_p = \frac{(E_p \times R_p \times P_p) \times \left(\frac{C_p}{ADP}\right)}{RR_p} \qquad (4.23)\ \ \Delta C_p = \frac{(E_p \times R_p \times P_p) \times \left(\frac{C_p}{POL}\right)}{RR_p}$$

The Risk Adjustment, *R*. The first adjustment factor is designed to modify E to account for how particular policy proposals may be designed for offenders with different risk-for-reoffense probabilities. For example, a policy change might be focused on early release from prison policies for lower-risk offenders.

The basic elasticity, E, was estimated from research studies that measure all offenders that make up the whole criminal population in question. If the models had been able to use "lower-risk" factor instead of total in the estimations, then E would have been different. The multiplicative adjustment factor, R, provides a way to model this likely result. We currently do not adjust our policing elasticities with a risk factor adjustment.

Washington State uses an actuarial-based risk assessment that predicts the probability of recidivism. This assessment is used in Washington to classify offenders in prison, in terms of recidivism risk, as lower risk, moderate risk, higher risk for non-violent recidivism, or higher risk for violent recidivism.[91] From the recidivism rates for all offenders and for those same offenders separated by risk levels, we compute simple ratios of recidivism rates. The ratios indicate the relative likelihood of recidivism for different risk levels, compared to all offenders as a group. These ratios are then used as the risk adjustment multipliers, R, in equations 4.20-4.23. Since there is risk around these risk adjustment multipliers, we use a triangular probability density distribution for the Monte Carlo simulation with minimum and maximum multiplicative values to account for between-group variation. The minimum and maximum parameters were estimated by examining the variation in cohort-to-cohort recidivism rates. We use the ratio relative to all offenders as illustrated in Exhibit 4.32 as the mean value and examine cohort-to-cohort variation to set the minimum and maximum values.

Exhibit 4.32

Three–Year Recidivism Rates of Offenders Released from Prison in Washington State, Fiscal Years 2002 to 2004

		Recidivism for a violent felony offense		Recidivism for a property felony offense	
Risk for re-offense category	**Number of offenders**	**Recidivism rate**	**Ratio: relative to all offenders**	**Recidivism rate**	**Ratio: relative to all offenders**
All offenders	14,459	12.8%	1.00	16.2%	1.00
Lower risk	2,018	3.6%	0.28	2.7%	0.16
Moderate risk	2,743	8.1%	0.63	9.3%	0.57
High risk, non-violent	5,167	9.3%	0.72	22.2%	1.37
High risk, violent	4,531	23.9%	1.86	19.6%	1.21

Note: Recidivism is defined as a new felony reconviction in the state of Washington within three years of release from prison, where the most serious conviction is either for a violent or property offense. For the purposes of Exhibit 4.32, other offenses, such a drug offenses, are not included in this definition.

The Policy Adjustment, *P*. Equations 4.20, 4.21, 4.22, and 4.23 implement a second multiplicative adjustment, P, to account for differences in the effectiveness of policies. Certain changes in prison term or policing strategies have evidence that indicates that these policies different from the general strategy

The Incarceration Policy Adjustment. There are two ways policies can affect total incarceration ADP: the probability of going to prison given a conviction and the length of stay given a prison sentence. The first factor implies punishment certainty while the second more closely reflects punishment severity. These two factors are likely to have different effects on crime, yet the overall elasticity, E, estimated with current research using total ADP, is unable to distinguish the separate effects. Therefore, equations 4.20 and 4.21 implement a second multiplicative adjustment, P, to account at least partially for this limitation in the current state of incarceration research. Without adjustment, simply using E to estimate how a change in prison length of stay affects crime would most likely over-estimate the effect.

Nagin (2013) and Durlauf & Nagin (2010) have found that changing length of stay is likely to have a smaller effect than changing the probability of punishment, we developed a procedure to provide a plausible adjustment to the overall prison-crime elasticity measured with the studies we include in the meta-analytic results displayed in Exhibit 4.33.[92] One of the steps of this procedure was to conduct a meta-analysis on the effect of the length of stay on crime. These results are below in Exhibit 4.33.

---

[91] Barnoski & Drake, (2007).

[92] Nagin, D. (2013). Deterrence in the twenty-first century: A review of the evidence. Crime and Justice: A Review of Research. Chicago, IL: University of Chicago Press.

Exhibit 4.33

Meta-Analytic Results: Prison Length of Stay on Recidivism

Topic	Dependent variable	Elasticity	Standard error	Number of studies
Prison LOS (a one month increase)	Crime	-0.010	0.009	9

Notes: All results are from random effects meta-analyses estimated with the methods described in Chapter 2.

To adjust the overall prison crime elasticity for length of stay policies, we implement the computational procedure displayed in Exhibit 4.34. To inform how length of stay policies affect current crime levels through incapacitation, we use our meta-analytic results measuring how length of stay affects the future recidivism rates of specific offenders display in Exhibit 4.33. If the effect of prison ADP on crime is primarily incapacitation rather than general deterrence, then studies of the effect of prison length of stay on the future recidivism rate of specific offenders provides useful estimates of how current crime levels change when length of stay changes. We estimate an elasticity metric for the literature estimating how prison length of stay affects the recidivism rate of specific offenders. From 1986 to 2009 in the US, prison length of stay increased by about four months, or about 17%, according to the US Department of Justice. We estimate that the 17% increase in length of stay resulted in roughly a 2% decrease in recidivism rates, as described computationally in Exhibit 4.34. This produces an elasticity of -0.202. Since the elasticity for total UCR crime from our meta-analysis reported in Exhibit 4.31 is -0.26, a simple policy multiplier to use to analyze length of stay policy changes with equations 4.20 and 4.21 is 0.776 (-0.202 / -0.26). Thus, when using the equations to analyze sentencing options that affect the length of prison stay on current crime levels, we use a mean multiplicative value of 0.776 to modify the overall elasticities reported in Exhibit 4.31 that measure both the probability or prison as well as the length of incarceration. The adjustment is rather crude (if data allowed, it would be better to estimate separate effects for violent and property crimes), but it does provide a first order approximation that is likely to be closer than simply using E as the effect. Since there is risk and uncertainty around this estimate, in Monte Carlo simulation we model a triangular probability density distribution with lower and higher values in addition to the modal value of 0.776.

Exhibit 4.34

Calculation of WSIPP Policy Adjustment Multiplier for Changes in Average Daily Prison Population Obtained by Changing the Length of Stay (rather than the probability of incarceration)

Step	Total crime
(1) Number of months change in prison length of stay, US, 1986 to 2009[1]	+4
(2) Percent change in length of stay[1]	+16.67%
(3) Effect size for change in recidivism, per month of prison length of stay[2]	-0.0102
Standard error[2]	0.09
(4) Effect size for observed change in length of stay[3]	-0.0408
(5) Base recidivism rate[4]	50%
(6) Recidivism rate after change in length of stay[5]	49%
(7) Percent change in recidivism rates[6]	-3.36%
(8) Elasticity: percent change in recidivism rate per percent change in length of stay[7]	-0.202
(9) Overall Prison/Crime elasticity[8]	-0.26
(10) Policy multiplier[9]	0.776

Notes:

1) Bureau of Justice Statistics, US Department of Justice, National Corrections Reporting Program, First Releases from State Prison, annual reports from 1986 to 2009. The mean length of stay increased from 24 to 28 months between 1986 and 2009.

2) Calculated from our meta-analysis of the effect of a one month increase in incarceration length of stay of criminal recidivism. Results are displayed in Exhibit 4.33.

3) We assume a linear effect size and multiply the effect size from step (3) times the number of months change from step (1).

4) This is roughly the long-term (15-year) recidivism rate of adults released from prison in Washington State, where recidivism is defined as a reconviction for a felony offense in Washington.

5) The recidivism rate after applying the Dcox effect size from step (4) to the base recidivism rate from step (5).

6) Step (6), divided by Step (5), minus one.

7) Step (7), divided by Step (2).

8) From Exhibit 4.31, the simultaneity adjusted elasticity for overall UCR crime.

9) Step (8), divided by Step (9).

The Policing Policy Adjustment. A growing body of research indicates that the way in which police are deployed in the community has a significant effect of crime rates. For example, Nagin's (2013) review of the literature found that "hot spots" and "pulling levers" policing deployment strategies have been shown to produce larger effects than traditional deployment strategies, while rapid response or thorough investigation strategies do not increase the effectiveness of policing on crime.[93] Thus, specific deployment policies are likely to have differential effects on crime, yet the overall elasticity, E, estimated with current research using total policing levels, is unable to distinguish additional effects. Therefore, equations 4.21 and 4.23 implement a policy adjustment, P, to account at least partially for this limitation in the current state of policing research.
For police elasticities, we adjust for the policing strategy being used, based on evidence that certain police strategies differ from average police deployment.

The steps we use to estimate a policing policy adjustment multiplier are listed in Exhibit 4.35 and follow this computational process:

$$(4.24)\quad PM_t = \frac{ME_t + \frac{(HSES_t \times SD_t \times \overline{POP})}{\overline{\overline{POL}}}}{ME_t}$$

We begin by computing the average marginal effect, ME, for crime type t, from our meta-analyses of the policing literature, described above. We then use the meta-analyzed effect size for hot spots policing, $HSES$, for crime type t, reported in the meta-analysis by Braga, et al. (2012).[94] The effect size measures, at the policing jurisdiction level, the effect of hot spots policing, in standard deviation units of crime, compared to non-hot spots jurisdictions. We use Washington State jurisdiction-level UCR data for 2011 in Washington's cities and county sheriff's offices for mean crime rates and the associated standard deviation in jurisdiction-level crime rates, SD, for crime type t. From the UCR data, we also include mean policing levels per jurisdiction, POL, and mean population per jurisdiction, POP. The resulting policy level multiplier estimates the degree to which policing following a hot spots deployment approach increases policing effectiveness relative to average effects, E. For example, a policy multiplier of 1.11 would indicate that hot spots deployed police are, on average, 11% more effective that police deployed with a routine strategy. We estimate an error term for the policy multiplier by running a Monte Carlo simulation, using the standard error from the Braga et al., (2012) meta-analysis.

[93] Nagin, (2013).
[94] Braga, A., Papachristos, A., & Hureau, D. (2012). *Hot spots policing effects on crime*. Campbell Systematic Reviews, 8.

Exhibit 4.35

Calculation of WSIPP Policy Adjustment Multiplier for Hot Spots Police Deployment

Step	Violent crime	Property crime
(1) Marginal effect of a police officer deployed with an average strategy, on annual UCR crime[1]	-1.89	-4.48
(2) Effect size of "Hot Spots" policing, compared to traditional deployment, jurisdiction level[2]	-0.175	-0.084
Standard error of the effect size	0.058	0.048
(3) Mean per-capita UCR crime rate in Washington policing jurisdictions[3]	0.00215	0.03147
Standard deviation in per capita crime rates	0.00177	0.01986
(4) Change in mean jurisdictional per-capita crime rate from hot spots deployment[4]	-0.00031	-0.00167
(5) Change in mean jurisdictional crimes from hot spots deployment[5]	-9.253	-49.794
(6) Change in crimes per officer from hot spots deployment[6]	-0.237	-1.278
(7) Mean Policy Adjustment Multiplier[7]	1.13	1.29
Washington State statistics		
Mean number of commissioned police officers per jurisdiction[8]	38.97	
Average population per jurisdiction[8]	28,852	

Notes:

1) Marginal effect (E*C/POL) calculated with an elasticity, E, times the current statewide level of violent or property UCR crimes, C, divided by the current statewide level of commissioned police officers. The elasticity, E, measures the average officer deployed in an average practice manner. The elasticities for the WSIPP analysis are reported in Exhibit 4.31.
2) From Table 10.4 of the meta-analysis by Braga, et al. (2012). Standard errors calculated from the confidence intervals reported in their Table 10.4.
3) Calculated from all reporting city and county sheriff's offices in Washington UCR data for 2011, with data reported on the website of the FBI.
4) The effect size from Braga, et al., (2012) times the standard deviation in crime rates for Washington jurisdictions.
5) The factor in footnote 4, times the average population per Washington policing jurisdiction, reported in this table.
6) Change in crimes per jurisdiction, divided by the mean number of officers per jurisdiction, reported in this table.
7) The sum of the marginal effect per officer (note one), plus the change in crimes per officer due to hot spots (note 6), divided by the marginal effect per officer.
8) Calculated for Washington police jurisdictions from UCR data and population data from the Washington State Office of Financial Management for 2011.

Estimating Large Changes in ADP or POL. Since the computation of marginal effects from equations 4.20, 4.21, 4.22, and 4.23 is designed for small unit changes in ADP or POL, and since the results will typically be used in practice to estimate the effects of larger policy changes in ADP or POL, the computation of the total marginal crime effect is estimated iteratively, one ADP or POL at a time. Equations 4.25, 4.26, 4.27, and 4.28 implement this iterative process for violent and property crime marginal effects. The equation sums the change in crimes for the (absolute value) of a total sentencing change or police change. For a policy that raises or lowers total prison ADP_T or total police levels POL_T, the change in crime by type, ΔC_V or ΔC_p, is calculated with the estimate of the adjusted elasticity for that type of crime, E times R times P, multiplied by the total crime of each type after each unit iteration of the total ADP or POL 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.

$$(4.25)\quad \Delta C_v = \frac{\sum_{a=1}^{|\Delta ADP_T|} (E_v \times R_v \times P_v) \times \frac{[C_{v(a)} + (\Delta C_{v(a-1)})]}{(ADP_T \pm a)}}{RR_v}$$

$$(4.26)\quad \Delta C_v = \frac{\sum_{a=1}^{|\Delta POL_T|} (E_v \times R_v \times P_v) \times \frac{[C_{v(a)} + (\Delta C_{v(a-1)})]}{(POL_T \pm a)}}{RR_v}$$

$$(4.27)\quad \Delta C_p = \frac{\sum_{a=1}^{|\Delta ADP_T|} (E_p \times R_p \times P_p) \times \frac{[C_{p(a)} + (\Delta C_{p(a-1)})]}{(ADP_T \pm a)}}{RR_p}$$

$$(4.28)\quad \Delta C_p = \frac{\sum_{a=1}^{|\Delta POL_T|} (E_p \times R_p \times P_p) \times \frac{[C_{p(a)} + (\Delta C_{p(a-1)})]}{(POL_T \pm a)}}{RR_p}$$

For example, for a policy that decreases prison ADP by 100 units, equations 4.25 and 4.27 are calculated 100 times, each time calculating the marginal crime effect after substituting a one unit reduction in ADP and the new level of the crime variable after the previous delta crime has been computed.

For a number of the benefit-cost calculations that follow, we are interested in total violent or property crime effects as described with equations 4.25, 4.26, 4.27, and 4.28. Total crime changes are used, for example, in computing the victim costs of crimes incurred or the victim benefits of crime avoided when policies change. For some calculations, however, we are only interested in computing the taxpayer costs of the criminal justice system and, hence for these calculations we are only interested in crimes reported to police. These reported-crime estimates, ΔRC_v and ΔRC_p are set using the following equations:

$$(4.29)\ \ \Delta RC_v = \Delta C_v \times RR_v$$

$$(4.30)\ \ \Delta RC_p = \Delta C_p \times RR_p$$

Victim Costs or Benefits. The victim costs or benefits are estimated with the following equation:

$$(4.31)\ \ \Delta Victim\$ = \Delta C_v \times VictimPerUnit\$_v + \Delta C_p \times VictimPerUnit\$_p$$

The change in the total value of victim costs, $\Delta Victim\$$, is the sum of the change in the number of violent and property victimizations from equations 4.25-4.28, ΔC_v and ΔC_p times, respectively, the marginal victim cost per violent and property victimization, $VictimPerUnit\$_v$ and $VictimPerUnit\$_p$. In Monte Carlo simulation, a triangular probability density distribution is used to model uncertainty in the per unit victim costs.

Criminal Justice System Costs or Benefits. When crime is increased or reduced, taxpayers can expect to pay more or less, respectively, from the policy change. The calculation of these amounts are done for police expenses; court-related expenses including court staff, prosecutor and defender staff; jail sanction costs; prison costs; and community supervision costs for jail-based or prison-based sentences. The change in expenses for each part of the criminal justice system are calculated using the following equations:

$$(4.32)\ \ \Delta Police\$ = \Delta RC_v \times \frac{Arrest_v}{RC_v} \times PolicePerArrest\$_v + \Delta RC_p \times \frac{Arrest_p}{RC_p} \times PolicePerArrest\$_p$$

$$(4.33)\ \ \Delta Court\$ = \Delta RC_v \times \frac{Conviction_v}{RC_v} \times CourtPerConviction\$_v + \Delta RC_p \times \frac{Conviction_p}{RC_p} \times CourtPerConviction\$_p$$

$$(4.34)\ \ \Delta Jail\$ = \Delta RC_v \times \frac{JailLOS_v}{RC_v} \times JailPerYear\$_v + \Delta RC_p \times \frac{JailLOS_p}{RC_p} \times JailPerYear\$_p$$

$$(4.35)\ \ \Delta Prison\$ = \Delta RC_v \times \frac{PrisonLOS_v}{RC_v} \times PrisonPerYear\$_v + \Delta RC_p \times \frac{PrisonLOS_p}{RC_p} \times PrisonPerYear\$_p$$

$$(4.36)\ \ \Delta JailCS\$ = \Delta RC_v \times \frac{JailCSLOS_v}{RC_v} \times JailCSPerYear\$_v + \Delta RC_p \times \frac{JailSuperLOS_p}{RC_p} \times JailCSPerYear\$_p$$

$$(4.37)\ \ \Delta PrisonCS\$ = \Delta RC_v \times \frac{PrisonCSLOS_v}{RC_v} \times PrisonCSPerYear\$_v + \Delta RC_p \times \frac{PrisonCSLOS_p}{RC_p} \times PrisonCSPerYear\$_p$$

For each segment of the criminal justice system, the change in expenses is the sum the change in the number of reported violent and property victimizations from equations 4.29 and 4.30, ΔRC_v and ΔRC_p times, respectively, the probability that a reported crime uses resources in each criminal justice segment, times the marginal cost of that segment per violent and property victimization. For jail and prison length of stay and for the length of stay on community supervision for jail-based and post-prison-based segments, the parameters are conditional on the probability of a conviction given a reported crime. The per-unit costs are denominated in a common "base" year's dollars used for all monetary valuations in the benefit-cost analyses. In Monte Carlo simulation, a triangular probability density distribution is used to model uncertainty in the marginal per-unit criminal justice costs.

The total change in crime-related costs from equations 4.31 to 4.37 and measures the effect of a policy change on current crime related costs or benefits is measured with the following equation:

$$(4.38)\ \Delta CurrentCrime\$ = \Delta Victim\$ + \Delta Police\$ + \Delta Court\$ + \Delta Jail\$ + \Delta Prison\$ + \Delta JailCS\$ + \Delta PrisonCS\$$$

4.3 Valuation of Child Abuse and Neglect Outcomes

WSIPP'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 Documentation describes WSIPP'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?[95] 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; WSIPP's model uses an incidence-based approach.

This component of WSIPP'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. WSIPP'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 (including the higher risk of death experienced by CAN victims). The third component is made up of other long-term costs that are causally linked to the incidence of CAN; these linkages are described in Section 4.3d and further detailed in the Appendix. 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; we are unable to monetize those effects at this time. [96]

4.3a CAN Prevalence and Cost 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. In addition, there are several other input parameters

---

[95] See for example, Wang, C.T. & Holton, J. (2007). *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

[96] 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.

used in the CAN model that are general to WSIPP's overall benefit-cost model; these are discussed elsewhere in this Chapter. In the following sections, the sources for the parameters and the computational routines are described.

Exhibits 4.36, 4.37, and 4.38 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.

WSIPP'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 4.36 displays the inputs, for age one to 18:

- the cumulative prevalence of CAN for general, low-income, and indicated populations;
- the annual likelihood of out-of-home placement for those with CAN for general and indicated populations;
- the cumulative likelihood of out-of-home placement for the imminent risk and SED populations.

Exhibit 4.36

Prevalence of CAN and OOHP by Population

	General population			Indicated (CWS-involved) population		Special populations	
	Abuse & Neglect: cumulative prevalence	Abuse & Neglect: cumulative prevalence (low-income)	Out-of-home placement of those with CAN	Abuse & Neglect: recurrence for maltreated children	Out-of-home place: of those with CAN	Children at imminent risk of removal	Children with serious emotional disturbance (SED)
	Rate of first substantiation, cumulative	Rate of first substantiation, cumulative, for low income population	CAN-referred population: age of 1st placement	Recurrent substantiation by follow up year, cumulative	Removal in each follow-up year	Placement by follow-up year, cumulative	Placement by follow-up year, cumulative
Age	Percent	Percent	Percent	Percent	Percent	Percent	Percent
1	0.0212	0.0451	0.3439	0.2124	0.3431	0.4911	0.3543
2	0.0302	0.0635	0.1303	0.3275	0.1984	0.5682	0.4076
3	0.0389	0.0810	0.1127	0.3949	0.1683	0.6133	0.4388
4	0.0469	0.0968	0.1025	0.4427	0.1508	0.6453	0.4609
5	0.0544	0.1113	0.0952	0.4797	0.1383	0.6701	0.4781
6	0.0615	0.1247	0.0896	0.5100	0.1286	0.6903	0.4921
7	0.0681	0.1371	0.0849	0.5356	0.1207	0.7075	0.5039
8	0.0743	0.1486	0.0811	0.5578	0.1140	0.7223	0.5142
9	0.0800	0.1590	0.0777	0.5774	0.1082	0.7354	0.5233
10	0.0853	0.1687	0.0747	0.5949	0.1031	0.7471	0.5314
11	0.0903	0.1776	0.0720	0.6107	0.0985	0.7577	0.5387
12	0.0949	0.1858	0.0696	0.6251	0.0944	0.7674	0.5454
13	0.0996	0.1939	0.0674	0.6384	0.0906	0.7763	0.5515
14	0.1042	0.2020	0.0654	0.6507	0.0872	0.7846	0.5572
15	0.1088	0.2098	0.0635	0.6622	0.0840	0.7922	0.5625
16	0.1133	0.2175	0.0618	0.6729	0.0810	0.7994	0.5675
17	0.1171	0.2239	0.0601	0.6830	0.0782	0.8062	0.5722
18	0.1195	0.2279	0.0586	0.6925	0.0755	0.8125	0.5766

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.[97] 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 victims[98] 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 one to age 18, the implied lifetime prevalence rate of child abuse or neglect for the general population of children is estimated to be 11.9%. The cumulative prevalence for CAN by age, after repeat cases are accounted for, is displayed in Exhibit 4.36.

To test the reasonableness of this estimate, we use a second approach to calculate the lifetime prevalence. We gather other research studies that examine this question with longitudinal cohort data. Exhibit 4.37 summarizes these estimates. The studies measure 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% lifetime prevalence of child abuse, slightly lower than, but similar to the estimate described in the first method above.

Exhibit 4.37

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 cases
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 group
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

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 those with low 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 examined this question with lower-risk control groups (see Exhibit 4.38).[99]

---

[107] Administration on Children, Youth and Families, (2011). *Child Maltreatment 2011* Table 3-4. Retrieved August 1, 2013, from http://www.acf.hhs.gov/sites/default/files/cb/cm11.pdf.

[98] Ibid., Table 3-13.

[99] 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 4.38

Odds Ratios for Child Abuse and Neglect: High-Risk Populations

Study	Number of participants in study	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 also displayed in Exhibit 4.36; we use child welfare history data from Washington State to estimate, of those children who receive one accepted referral, the proportion who subsequently receive another accepted referral over time.[100] We analyze the proportion of children who have experienced a recurrence of abuse or neglect, from one year out to 12 years. We then plot a logarithmic curve with those data to predict the likelihood of a recurrence from up to 17 years after the initial incident.

Exhibit 4.36 also displays the base rates of out-of-home placement for various populations. For the general population, we calculate the probability of out-of-home placement at each age, given a child has an accepted CAN referral, based on a WSIPP analysis of Washington State child welfare data.[101] To compute the base likelihood of out-of-home placement for a prevention population, we multiply the likelihood of a substantiated CAN case at each age (derived from NCANDS data as described above) by the ratio of Washington-reported accepted referrals to estimated CAN cases,[102] then by the likelihood of out-of-home placement given CAN at each age.

For the population of children already in the child welfare system, we computed the likelihood, for each year following a second accepted referral (regardless of their age at first or second accepted referral), that a child would be removed from home. For children deemed at "imminent risk" of placement, a WSIPP analysis determined the risk of out-of-home placement for these children was much higher than in the indicated population (from the studies we included, about 25% of children at "imminent risk" of placement had been removed from home in the first three months; this number grew to nearly 50% by one year).[103] Our analysis resulted in a unique predicted base rate of out-of-home placement for the "imminent risk" population. The last column in Exhibit 4.36 shows the cumulative likelihood over time 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.[104]

---

[100] WSIPP analysis of DSHS CAMIS data for FY 1998 and FY 2000 birth cohorts.

[101] Using data from DSHS CAMIS for children born between July 1, 1997 and July 1, 2008, we examined the subset of children who had at least one accepted referral at some point in their childhood (in our analysis, accepted referrals act as a proxy for substantiated CAN cases; later in the analysis we compute the ratio of accepted referrals to our estimate of substantiated CAN cases as an adjustment). We computed the proportion of children who were removed at some point subsequent to that accepted referral, by age of first accepted referral.

[102] To compute this ratio, we use data from DSHS CAMIS for children born between July 1, 1997 and July 1, 2008 to determine what proportion had at least one accepted referral by age 11. We then divide this proportion by our estimated cumulative proportion of substantiated CAN in the general population by age 11 (see Exhibit 4.36).

[103] WSIPP analysis of two evaluations of the HOMEBUILDERS® model of intensive family preservation services, which serve youth at "imminent risk" of placement and report cumulative likelihood of out-of-home placement at different periods of time. We plotted the likelihood of placement by follow-up period and fit a logarithmic curve to the point-in-time estimates, projecting rates of removal for up to 17 years.

[104] We calculated the cumulative percent from two studies of Multisystemic Therapy for children with SED that followed children over more than one year. We used the data from four points in time to plot a logarithmic curve from which we projected rates of placement for up to 17 years.

Estimated per-child child welfare system costs are displayed in Exhibit 4.39. 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.

In addition, we also estimate the cost of placing children with serious emotional disturbance (SED);[105] although these children are not placed for reasons of abuse and neglect, but rather mental health problems, the programs which aim to avoid placements for this population are often provided within the child welfare system.

Exhibit 4.39
The Estimated Average Public Cost of a Child Protective Service Case Accepted for Investigation, State of Washington (in 2014 Dollars)

	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	37,992[1]	**100%**	**$696[2]**	**2011**	**$727**
Police involvement	6,345[3]	16.7%	$670[4]	2009	$122
Juvenile court dependency case involvement	4,864[5]	12.8%	$3,373[6]	2007	$484
Child welfare services					
Percentage of protective custody placements that are CPS cases	75.27%[7]				
Protective custody (foster care)	5,589[8]	11.1%	$34,623[9]	2012	$3,931
In-home services (not out-of-home placement)	37,992[10]	2011	$462	2011	$483
Adoption	790[11]	2.1%	$79,094[12]	2012	$1,686
Juvenile court termination case involvement	1,705[13]	4.5%	$3,906[6]	2007	$196
TOTAL: Expected present value cost of an accepted CPS case (in 2014 dollars)					**$7,630**
Rate of decay of system costs over time					**-0.53**
Addendum: Expected present value cost of an out-of-home placement, conditional on an out-of-home placement					**$48,300**
Addendum: Expected present value cost of an out-of-home placement, for a child with serious emotional disturbance (SED)					**$9.026**

Notes:
1) Washington State DSHS Children's Administration, 2011 Year in Review, available at: http://www.dshs.wa.gov/pdf/ca/year-in-review2011.pdf
2) Washington State DSHS Research and Data Analysis Client Data for FY2011. Total expenditures for "Child Protective Services case management", divided by total accepted referrals.
3) Percentage of referrals from police sources, all states, applied to total accepted referrals. From Administration on Children, Youth and Families (2011) Child Maltreatment 2011, Table 2-C, available at: http://www.acf.hhs.gov/sites/default/files/cb/cm11.pdf.
4) Marginal operating cost of an arrest for a misdemeanor from WSIPP crime model.
5) Washington State Office of the Administrator of the Courts, 2012, Juvenile dependency filings. Report available at http://www.courts.wa.gov/caseload/?fa=caseload.showReport&level=s&freq=a&tab=juvDep&fileID=jdpfilyr.
6) Based on average number of hearings per case (see Miller, M. (2004). *How do court continuances influence the time children spend in foster care* (Doc. No. 04-03-3901). Olympia: Washington State Institute for Public Policy.) multiplied by WSIPP analysis of average cost per hearing (based on projected length in hours, and the hourly wages for the people estimated to be involved in each hearing).
7) Based on WSIPP analysis of DSHS Children's Administration data.
8) AFCARS 2011, Children in Foster Care (entered care): http://cwoutcomes.acf.hhs.gov/data/downloads/pdfs/washington.pdf
9) Calculated based on DSHS Children's Administration projected per-capita costs for FY2013. We recognize that there are additional costs of out-of-home care for children placed with relatives, such as child-only TANF payments. We are unable to estimate these costs at this time, but plan to do so in the future.
10) DSHS Children's Administration EMIS reporting system; unduplicated counts of children served are unreported; therefore, we summed FY11 total costs for in-home services and divided by total accepted referrals for a cautious per-child estimate.
11) WSIPP estimate of new adoption cases each year, from FY2008 DSHS Children's Administration data.
12) WSIPP calculation of total adoption support per case, estimated from FY2012 Children's Administration data.
13) Washington State Office of the Administrator of the Courts, 2012, Juvenile termination filings. Report available at http://www.courts.wa.gov/caseload/?fa=caseload.showReport&level=s&freq=a&tab=juvDep&fileID=jdpfilyr

[105] The cost of out-of-home placement for SED children is based on a WSIPP analysis of Washington State data, taking into account the cost for Behavioral Rehabilitation Services (BRS—residential treatment for children) and the average length of stay in such treatment. Cost data was derived from the DSHS Children's Administration EMIS reporting system (average monthly per-child ongoing placement services costs for FY11), and length of stay was estimated from DSHS CAMIS data for children removed from home for behavior, drug, or alcohol problems between January 1, 1999 and January 1, 2005.

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 4.40.[106] These estimated totals are life cycle expected value costs per CAN crime; we use the "decay" parameter for victim costs above to "spread out" those costs over a child's life.

Exhibit 4.40
Medical, Mental Health, and Quality of Life Costs
per Victim of Child Abuse and Neglect, 1993 Dollars

	Medical and mental health costs(1) (1)	Quality of life costs(1) (2)	Number of victims(3) (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**	
Rate of decay of victim costs over time	**-.10**		

Sources:
1) The source of the cost elements in this table is Miller et al., (2001).
2) *Ibid*., Table 1. We assumed 80% urban and 20% rural costs on the Miller et al. Table 1.
3) The source for the total US number of victims: Miller, T.R., Cohen, M.A., & Wiersema. B. (1996). *Victim costs and consequences: A new look*. 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) WSIPP assumptions.

The final parameters in Exhibits 4.39 and 4.40 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 4.39, we estimate the amount of system-related costs we would expect to be incurred within the first two years of a typical CAN case (78%). Using that figure, we calculate a rate of "decay," such that for each year after the beginning of a case, the amount of cost decayed by -0.53. That means, in the first year, 53% of the total expected costs are incurred; by the end of the second year, 78% have been incurred; 90% 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., often do not continue past the age of 17.

We also estimate 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 is -0.10, 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.

Sources of CAN and OoHP costs. The parameters described in Exhibit 4.41 allow users to input the proportion of child welfare funding from state, local, and federal sources.

[106] Miller, T.R., Fisher, D.A., & Cohen, M.A. (2001). Costs of juvenile violence: Policy implications. *Pediatrics, 107*(1).

Exhibit 4.41
Proportion of CAN and OoHP Costs by Source

	State	Local	Federal
CPS response[1]	0.625	0.000	0.375
Police involvement[2]	0.150	0.850	0.000
Juvenile court (dependency)[3]	0.510	0.490	0.000
Protective custody (foster care)[1]	0.625	0.375	0.000
In-home services[1]	0.625	0.375	0.000
Adoption[4]	0.500	0.000	0.500
Juvenile court (termination)[3]	0.440	0.560	0.000
Out-of-home placement for children with SED[4]	0.500	0.000	0.500
Victimization (taxpayer) costs[5]	0.500	0.000	0.500

[1] For the 75% of kids who are Title IV-E eligible, we apply the Washington State FMAP rate from Federal Register /Vol. 75, No. 217 /November 10, 2010 /Notices 69083, accessed from: http://aspe.hhs.gov/health/fmap12.pdf. For the 25% of non-eligible children, we assume the state pays 100%.
[2] Justice Expenditure and Employment Extracts, 2010 - Preliminary, Tracey Kyckelhahn, Ph.D., Tara Martin, BJS Intern, July 1, 2013. NCJ 242544, Table 4: Justice system expenditure by character, state and type of government, fiscal 2010, Link: http://www.bjs.gov/index.cfm?ty=pbdetail&iid=4679. Direct current Police Protection expenditures for state and local governments for Washington State.
[3] WSIPP analysis of staff present at juvenile hearings; assume state pays 100% of Assistant Attorney General and social worker salaries, 50% of judicial officer salaries. Other staff are assumed to be fully funded by the local government.
[4] Department of Health and Human Services, *75*(217) Fed. Reg. 69083 (proposed Nov. 10, 2010), accessed from: http://aspe.hhs.gov/health/fmap12.pdf.
[5] We assume that victim costs to taxpayers will be in form of health and mental health treatment; with 50/50 FMAP split.

4.3b Deaths Attributed to CAN

Children who are victims of CAN have a higher risk of death than children who are not. Data collected by the Children's Bureau at the federal Administration for Children and Families give the number of children who die each year as a result of abuse or neglect.[107] We use these numbers to compute the likelihood of death by age for CAN victims (see Exhibit 4.42). We assume that interventions that reduce the likelihood of CAN also reduce the risk of death by CAN, so we apply the risk of death by CAN at each age post-treatment to the amount of change we expect an intervention to cause by age, then multiply by the value of a statistical life (as described in Section 4.11d) for each age.

Exhibit 4.42
CAN attributed deaths by age, United States, 2013

Age group	Years in age group	CAN attributed deaths in US	All deaths in US	US population
Less than 1 year	1	707	23,440	3,941,783
Age 1-3	3	524	3,423	11,934,615
Age 4-7	4	178	2,153	16,363,731
Age 8-11	4	53	1,802	16,327,716
Age 12-15	4	40	3,076	16,668,723
Age 16-17	2	15	3,193	8,349,304

[107] Children's Bureau (2015). Child abuse and neglect fatalities 2013: Statistics and interventions, accessed from https://www.childwelfare.gov/pubs/factsheets/fatality.

4.3c Linkages: CAN and Other Outcomes

WSIPP's benefit-cost model monetizes improvements in CAN, in part, with linkages between CAN and other outcomes for which a monetary value can be estimated. For example, credible research shows a causal link between the incidence of CAN and subsequent criminal behavior of the victimized youth when he or she is older. The parameters for these linkages are obtained by a meta-analytic review of relevant research literature. For example, we estimate the relationship between CAN and later participation in crime 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 parameters are entered into the benefit-cost model and used when performing a Monte Carlo simulation. The linkages in the current WSIPP model are listed in the Appendix.

The studies that allow us to estimate causal links between child abuse and neglect and other, longer-term outcomes are most often based on the relationship between any CAN and some later consequence. While it is clear that there are consequences caused by one or more experiences of CAN (compared to zero experiences of CAN), there is not enough evidence for us to judge whether those relationships hold true for children who have already experienced CAN (and for whom we estimate some reduction in further CAN).To be cautious, we cut the magnitude of each estimated link in half when estimating benefits for CAN reduction for intervention populations (children who have already experienced some amount of CAN).

To model the human capital outcomes affecting labor market earnings via CAN, we follow the same procedures described in depth in Section 4.4d. In our examination of the research literature, we found a strong effect of CAN on the probability of employment as an adult, but no evidence to suggest that the earnings of CAN victims if employed would be any different than non-victims. For intervention populations, we apply our assumption about the reduced magnitude to this effect size. We then fit distributions of expected earnings given CAN using the methodology described in Section 4.4d. Exhibit 4.43 shows the parameters for the fitted distributions that reflect the changes in earnings.

Exhibit 4.43

Labor Market Parameters for CAN Morbidity and Mortality

	Gain in labor market earnings for prevention of CAN vs. CAN experiences	Gain in labor market earnings for CAN intervention vs. further CAN experiences
Distribution type	LogNormal	LogNormal
Mean	-1.0761	-1.2863
Std Dev.	0.1552	0.1548
Shift	0.7777	0.7767

4.4 Valuation of Alcohol, Illicit Drug, and Regular Tobacco Use Outcomes

WSIPP'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 WSIPP's model divides drugs into a) cannabis, b) opioids, and c) all other illicit drugs.[108] Analysts sometimes abbreviate alcohol, tobacco, and other drugs with the acronym ATOD. This section of the Technical Documentation describes WSIPP's current procedures to estimate the monetary benefits of program-induced changes in ATOD. For WSIPP's benefit-cost model, an alcohol and illicit drug disorder reflects either abuse or dependency as defined by the Diagnostic and Statistical Manual (DSM) of the American Psychiatric Association. Regular smoking is defined as daily smoking.

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

In general, analysts construct two types of studies to estimate the costs of ATOD: "prevalence-based" studies and "incidence-based" studies.[109] 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.

WSIPP'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 attempt to estimate these values.[110] 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.[111] 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.[112] These prevalence-based total cost studies can be valuable, but they are not designed to evaluate future marginal benefits and marginal costs of specific public policy options.

The purpose of WSIPP'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% 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 substance use disorders. Depending on each particular substance, the following cost categories are included in WSIPP'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.
- Medical costs for hospitalization, emergency department, and pharmaceuticals or total health care costs from ATOD morbidity or mortality, to the degree that these costs are caused by ATOD.
- Crime costs to taxpayers and victims, to the degree that crime is estimated to be caused by ATOD.
- Traffic collision costs, to the degree that collisions are estimated to be caused by ATOD (only used in the case of alcohol).
- Treatment costs of ATOD, to the extent that disordered users of ATOD utilize treatment.
- Value of a statistical life (VSL) estimates cost to society, net of labor market changes, applied to the change in mortality estimated to be caused by ATOD.

[109] 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.
[110] 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.
[111] 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.
[112] Wickizer, T.M., (2007). *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.

4.4a ATOD Epidemiological Parameters: Current Prevalence for Prevention and Intervention Programs

WSIPP's ATOD model begins by analyzing the epidemiology of each ATOD disorder or problem to produce estimates of the current 12-month prevalence of heavy and disordered alcohol use, disordered cannabis, opioid, and other illicit drug use, and regular tobacco smoking (we use the general phrase "ATOD disorder" to refer to any of these conditions). 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.

The ATOD model also provides the base methodology for computing the current prevalence of other health conditions, including depression, anxiety, ADHD, disruptive behavior disorders, serious mental illness, post-traumatic stress disorder, diabetes, and obesity.

The formulas presented here are used not only in the ATOD model, but also in the mental health and health care models. Later chapters describing methods for these topic areas will refer back to Section 4.4a.

Four parameters enter the model to enable an estimate of the current prevalence of ATOD, from age one 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, and
- Death (survival): the probability of death by age, after the age of treatment by a program.

Exhibit 4.45 displays the current parameters in WSIPP's model for the first three epidemiological factors, along with sources and notes. The death probability information is described in Section 4.4b.

For each ATOD disorder, or other health condition, the current prevalence among the general population is estimated using the following equation:

$$(4.39) \quad CPG_y = \left(\sum_{0=1}^{y} O_0 \times P_{(y-0+1)} \right) \times LTP \times S_y \times SF_a$$

The current prevalence probability at any year in a person's life, CPy, is computed with information on the age-of-onset probability, O, from prior ages to the current age of the person, multiplied by the persistence probability, P, of remaining in the condition at each onset age until the person is the current age, multiplied by the lifetime probability of ever having the condition, LTP, multiplied by the probability of any-cause survival at each age, S_y, multiplied by the probability of condition-related survival in each age group, SF_a, following treatment by a program.

For each ATOD disorder or health condition, the exogenous age-of-onset probability distribution for ages one to 100, O, is a density distribution and is estimated with information from the sources shown in Exhibit 4.45.

$$(4.40) \quad 1 = \sum_{y=1}^{100} O_y$$

Also, for each ATOD disorder or health condition, the exogenous persistence distribution for ages after onset, P, is computed from the sources shown in Exhibit 4.45. The persistence distribution describes the probability, on average, of being in the condition each year following onset.

The probability of survival at any given age (all causes), 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 Section 4.11c of the Technical Documentation. To compute the current prevalence of a disorder over the entire life course, S_y is normalized to age one, as given by the following equation:

$$(4.41) \quad S_y = \frac{LTS_y}{LTS_1}$$

Because 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, as it is assumed that all treatment programs will be for those currently alive at time of treatment, as shown in the following equation:

$$(4.42)\quad S_y = \frac{LTS_y}{LTS_{tage}}$$

The final term in equation 4.39 is the reduced chance of survival due to the specific health condition, above and beyond what one may observe generally. For individuals in the general population, we compute estimates for each age group with the following equation:

$$(4.43)\quad SFG_a = \frac{1 - \left(\frac{CondD_a}{(Pop_a \times CP_a)} + \frac{PopD_a - CondD_a}{(Pop_a)}\right)}{\left(1 - \frac{PopD_a}{Pop_a}\right)}$$

In this equation 4.43, Pop_a is the total population in a state in each age group, CP_a is the average current prevalence in each age group, $PopD_a$ is the total number of deaths in a state in each age group, and $CondD_a$ is the deaths attributable to the ATOD disorder or other health condition in each age group.

Equation 4.39 describes the calculation of current prevalence for general (prevention) populations. For programs treating indicated populations, CPI_y the prevalence in all years following treatment is described using the following equation:

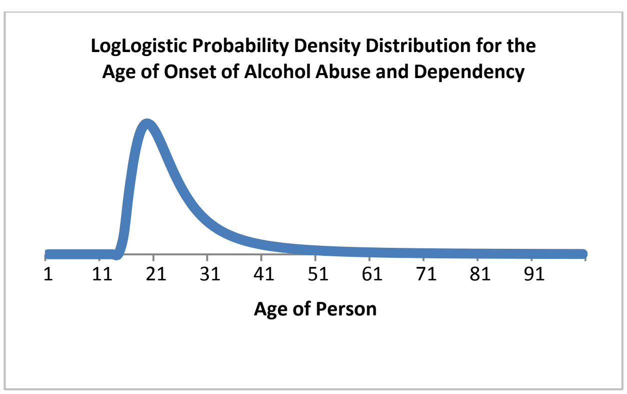
$$(4.44)\quad CPI_y = \frac{\sum_{0=1}^{tage} O_0 \times P_{(y-0+1)}}{\sum_{0=1}^{tage} O_0} \times S_y \times SFI_a$$

Finally, the survival factors for indicated populations by age group (SFI_a) can be calculated with the following equation:

$$(4.45)\quad SFI_a = (SFG_a \times CP_a) + (1 - CP_a)$$

Example. We provide an illustrative example of computing CPG_y in equation 4.39 for disordered alcohol use. Using the results from Hasin et al., we computed a probability density distribution for the age of onset of DSM alcohol disorders.[113] 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 4.44. The exhibit below plots the estimated distribution, where the sum of annual probabilities equals 1.0.

Exhibit 4.44



[113] 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.

Next, estimates of the persistence of the alcohol disorder, given onset, were computed for alcohol from a study by Lopez-Quintero, et al.[114] 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 4.45 and Exhibit 4.46 plots the results.

[114] 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 4.45

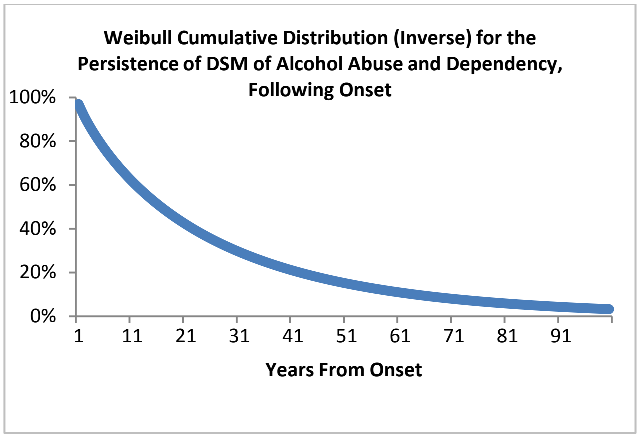
Input Parameters for the Epidemiology of Alcohol Disorders, Illicit Drug Disorders, and Regular Smoking[1]

	DSM alcohol disorder (a)	Heavy drinking (b)	DSM illicit drug disorder (cannabis) (c)	DSM illicit drug disorder (non cannabis) (d)	DSM illicit drug disorder (opioids) (e)	Regular tobacco smoking (f)
Percentage of population with lifetime DSM disorder, heavy drinking, or regular smoking	24.2%[2]	31.3%[7]	8.5%[8]	5.5%[8]	1.53%[10]	39.3%[13]
Age of onset:						
Type of distribution	Log-logistic[3]	Log-logistic[3]	Extreme value[9]	Extreme value[9]	Log-logistic[11]	Log-logistic[14]
Parameter 1	14.5776	14.5776	18.0348	18.0348	9.4332	4.5788
Parameter 2	8.0661	8.0661	3.6638	3.6638	8.344	12.647
Parameter 3	2.05	2.05	n/a	n/a	2.264	6.8346
Parameter 4	n/a	n/a	n/a	n/a	n/a	n/a
Remission of DSM disorder, given onset						
Type of distribution	Weibull[4]	Weibull[4]	Lognormal[4]	Lognormal[4]	Weibull[12]	Beta-general[4]
Parameter 1	0.5	0.5	1.7917	1.4741	0	0.5
Parameter 2	0.86728	0.86728	1.149	1.0985	0.74791	0.96399
Parameter 3	24.129	24.129	n/a	n/a	9.7642	2.0358
Parameter 4	n/a	n/a	n/a	n/a	n/a	0
Parameter 5	n/a	n/a	n/a	n/a	n/a	115.25
Percentage of general population consuming substance	67.2%[5]	67.2%[5]	11.4%[5]	8.4%[5]	3.9%[13]	27.8%[5]

Notes and sources

1) 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. Heavy drinking is defined by exceeding recommended maximum weekly or both daily and weekly drinking limits. All outcomes are estimated as dichotomous conditions.
2) Vergés, A., Littlefield, A.K., & Sher, K.J. (2011). Did lifetime rates of alcohol use disorders increase by 67% in 10 years? A comparison of NLAES and NESARC. Journal of Abnormal Psychology, 120(4), 868-77 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 (0.1817 and 0.3028). When the averaged lifetime value is entered into our model, the resulting current prevalence estimate from our model (0.077) is nearly identical to the average of the current prevalence estimates, reported by Vergés, from the two national surveys (0.079, the average of 0.0740 and 0.0846).
3) Hasin et al. (2007). 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. We investigated the onset of heavy drinking in a separate analysis, using Kalaydjian et al.'s (2009) analysis of NCS-R data (Kalaydjian, A., Swendsen, J., Chiu, W.T., Dierker, L., Degenhardt, L., Glantz, M., Merikangas, K.R., ... Kessler, R. (2009). Sociodemographic predictors of transitions across stages of alcohol use, disorders, and remission in the National Comorbidity Survey Replication. *Comprehensive Psychiatry, 50,* 4). We used the definition of the first time individuals reported experiencing a symptom of alcohol abuse; the onset curve was nearly identical to estimates of abuse and dependence onset from Hasin et al. (2007), so elected to use parameters derived from Hasin et al. (2007) for both disordered and heavy alcohol use.
4) Lopez-Quintero et al. (2011). 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. We were unable to estimate a separate curve for heavy drinking and therefore used the remission parameters from alcohol abuse or dependence for heavy drinking as well.
5) 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.
6) Estimated based on the ratio of lifetime to past-year alcohol abuse or dependence, reported in Vergés et al., (2011), Table 2. We used past year reported heavy drinking (see Note 1 above) from the NESARC (Chen, C.M., & National Institute on Alcohol Abuse and Alcoholism (US). (2006). *Alcohol use and alcohol use disorders in the United States: Main findings from the 2001-2002 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC)*. Bethesda, Md: National Institute on Alcohol and Alcoholism.) , multiplied by the ratio derived from Verges et al. (2011).
7) 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.
8) *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.
9) Analysis of Wave 1 NESARC data; we combined the variables HER12ABDEP, HERP12ABDEP, PAN12ABDEP, and PANP12ABDEP to identify the weighted proportion of individuals with reported abuse or dependence on heroin and/or opioid drugs in the past 12 months or prior to the past month.
10) Huang, B., Dawson, D.A., Stinson, F.S., Hasin, D.S., Ruan, W.J., Saha, T.D., Smith, S.M., ... Grant, B.F. (2006). Prevalence, correlates, and comorbidity of nonmedical prescription drug use and drug use disorders in the United States: Results of the National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of Clinical Psychiatry,* 67(7), 1062-73. From the Figure reported in the Huang paper, we computed a loglogistic probability density distribution for the age of onset of a DSM drug disorder, conditional on having a disorder. Although our definition of opioids includes heroin, the Huang paper does not, nor could we find any alternative estimates for heroin disorder onset.
11) Blanco, C., Secades-Villa, R., García-Rodríguez, O., Labrador-Mendez, M., Wang, S., & Schwartz, R. P. (2013). Probability and predictors of remission from lifetime prescription drug use disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Psychiatric Research,* 47(1), 42-9. We fitted a Weibull cumulative probability distribution to the remission information reported in the study, and then inverted to estimate persistence curves.
12) UNODC, *World Drug Report 2013* (United Nations publication, Sales No. E.13.XI.6), page 2: opioid use in North America.
13) 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.
14) 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.

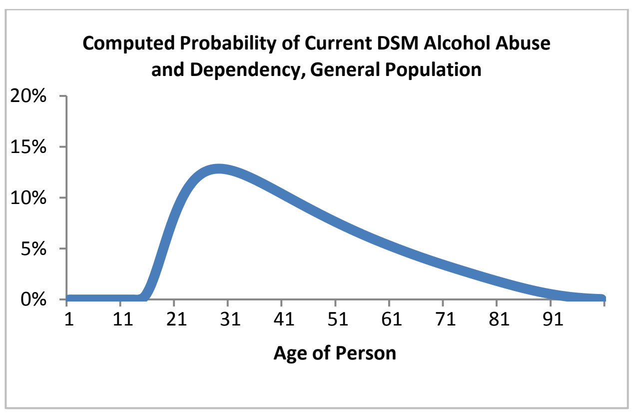
Exhibit 4.46



The persistence curve, after multiplying by the survival factor, by year, from the 2006 US 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% with sources shown in Exhibit 4.45, and after adjusting for survival from the 2006 US life table (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 4.39 and is plotted in Exhibit 4.47.

Exhibit 4.47



The same procedures just described for alcohol disorders are used for heavy alcohol use, disordered illicit drug use (non-cannabis), DSM cannabis use, DSM opioid use, and regular tobacco smoking, substituting the relevant parameters for the best-fitting distributions as shown in Exhibit 4.45. As noted, the estimates of the current prevalence of 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 disorder.

4.4b ATOD Attributable Deaths

WSIPP'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 4.48 for alcohol, 4.49 for tobacco, 4.50 for illicit drugs other than cannabis, and 4.51 for opioid drugs.

Alcohol. For alcohol-attributable deaths, the data source is the US Department of Health and Human Services, Centers for Disease Control (CDC). CDC estimates, for each state, the number of deaths attributable to alcohol causes.

Exhibit 4.48
Alcohol Attributable Deaths by Year, 2006-2010

Age group	Years in age group	Alcohol attributed deaths: chronic	Alcohol attributed deaths: acute	% of deaths attributable to DSM alcohol	% of deaths attributable to problem alcohol	All deaths in state	State population in age group
0-19	20	2	51	0.50	0.75	823	1,759,490
20-34	15	9	237	0.50	0.75	1,089	1,370,833
35-49	15	174	260	0.50	0.75	1,338	1,430,668
50-64	15	405	216	0.50	0.75	9,216	1,251,512
65-100	36	335	282	0.50	0.75	35,079	777,554

The estimates from CDC are available on-line via a software application called *Alcohol-Related Disease Impact (ARDI)*.[115] 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 cut points 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 4.48, for the years 2006-10. ARDI estimates deaths related entirely or partially due to particular causes of death. For the deaths partially caused by alcohol, we obtain 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 WSIPP's model focuses on DSM-level alcohol disorders and heavy drinking, a portion of the deaths caused by acute conditions could be from alcohol-involved events of someone who does not have a DSM-level condition and is not a habitually heavy drinker. Therefore, for acute deaths, the input screen provides for two parameters, by age group, to estimate the proportion of acute alcohol-related deaths where a DSM-alcohol disordered person was involved, and the proportion where heavy drinkers were likely involved.

To compute alcohol induced death rates for these age groups, we obtain 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 2006-10, the same years as the CDC/ARDI death estimates.

Tobacco Smoking. For smoking-attributable deaths, the data source is also the US Department of Health and Human Services, Center for Disease Control (CDC). CDC estimates, 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)*.[116] 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.

---

[115] Centers for Disease Control and Prevention website: https://apps.nccd.cdc.gov/ardi/HomePage.aspx
[116] Centers for Disease Control and Prevention website: http://apps.nccd.cdc.gov/sammec/

Exhibit 4.49

Smoking Attributable Deaths by Year, 2008

Age group	Years in age group	Smoking attributed deaths	All deaths in state	State population in age group
0-34	35	0	1,991	3,143,100
35-44	10	116	1,330	931,508
45-54	10	518	3,524	989,430
55-64	10	1217	5,864	768,070
65-74	10	1582	7,571	413,358
75-84	10	2262	12,368	251,045
85-100	16	1456	15,902	111,734

Illicit Drugs and Opioid Drugs. For illicit drug deaths, we use Washington State death data from CDC Wonder[117] for the years 2006 to 2010. For opioid drug deaths, we use data from the Washington State Department of Health Opioid File from 2012. We compute average annual drug attributable deaths in the age groups shown in Exhibit 4.50 for other illicit drugs and in Exhibit 4.51 for opioids.

Exhibit 4.50

Illicit Drug Attributable Deaths by Year, 2006-2010

Age group	Years in age group	Illicit drug attributed deaths	All deaths in state	State population in age group
0-14	15	0	604	1,295,338
15-19	5	22	220	464,152
20-24	5	55	354	470,333
25-34	10	159	735	900,500
35-44	10	226	1,338	929,838
45-54	10	329	3,428	984,350
55-64	10	153	5,787	767,993
65-74	10	34	7,500	414,866
75-84	10	14	12,146	251,414
85-100	16	8	15,432	111,273

[117] Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying Cause of Death 1999-2010 on CDC WONDER Online Database, released 2012. Data are from the Multiple Cause of Death Files, 1999-2010, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Accessed at http://wonder.cdc.gov/ucd-icd10.html on Jan 21, 2014.

Exhibit 4.51
Opioid Attributable Deaths by Year, 2012

Age group	Years in age group	Opioid attributed deaths	All deaths in state	State population in age group
0-14	15	1	632	1,309,139
15-19	5	11	190	449,500
20-24	5	50	352	467,031
25-34	10	125	810	946,195
35-44	10	117	1,216	905,468
45-54	10	203	3,324	966,058
55-64	10	137	6,437	880,718
65-74	10	28	8,422	512,730
75-84	10	9	11,965	257,808
85-100	16	5	16,708	123,123

For each ATOD, the death data are used to compute the probability of dying from ATOD in the general population, by age group, using the following equation:

$$(4.46) \quad AtodD_a = \big((Chronic_a + Acute_a \times AcutePct)/Pop_a\big)/Years_a$$

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. The value of the death is monetized with the value of a statistical life described in Section 4.11d.

4.4c Linkages: ATOD and Other Outcomes

WSIPP'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 parameters are entered into the benefit-cost model and used when performing a Monte Carlo simulation. The linkages in the current WSIPP model are listed in the Appendix.

4.4d Human Capital Outcomes Affecting Labor Market Earnings via ATOD-Caused Morbidity and Mortality

The WSIPP 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 Chapter 4.1, WSIPP's model uses national earnings data from the US 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, but do not currently have a disorder. 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 the following equation:

$$(4.47) \quad EarnC_y = \frac{EarnAll_y \times (1+EarnEscAll)^{y-tage} \times EarnBenAll \times (1+EarnBenEscAll)^{y-tage} \times StateAdj \times (IPD_{base}/IPD_{cps})}{\left((1+EarnGN) \times \left(1-\left(CP_y + \left(\sum_{o=1}^{y}(O_o \times LTP) - CP_y\right)\right)\right) + (1+EarnGF) \times \left(\sum_{o=1}^{y}(O_o \times LTP) - CP_y\right) + CP_y\right)}$$

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*, a Washington-specific earnings adjustment, StateAdj, and an adjustment to denominate the year of the CPS earnings data, *IPDcps*, with the year chosen for the overall analysis, *IPDbase*. These variables are described in Chapter 4.1.

The denominator uses the epidemiological variables described above: age of onset probabilities, *Oy*, lifetime prevalence rates, *LTP*, and current 12-month prevalence rates, *CPy*, 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). These relationships are derived from meta-analytic reviews of the relevant research literature.

For ATOD disorders, we meta-analyze 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. The Appendix displays the results of our meta-analysis of these two bodies of research for each ATOD disorder. Our meta-analytic procedures are described in Chapter 2.

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 2013 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 ATOD 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. The distribution parameters are shown in Exhibit 4.52. In the Monte Carlo analysis, we randomly draw probabilities as seeds for the modeled distribution. 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 parameters for both the *EarnGN* and the *EarnGF* variables.

Exhibit 4.52

Labor Market Earnings Parameters for ATOD Disorders

		DSM alcohol disorder	Heavy drinking	DSM illicit drug disorder (cannabis)	DSM illicit drug disorder (non cannabis)	DSM illicit drug disorder (opioids)	Regular tobacco smoking
Gain in labor market earnings for never used vs. current disordered users, probability density distribution parameters	Distribution Type	Gamma	Gamma	Gamma	Gamma	Gamma	Normal
	Alpha/Mean	39.608	53.157	47.014	47.014	47.014	1.9276
	Beta/Std Dev.	0.01255	0.00890	0.0052	0.0052	0.0052	0.03059
	Shift	0.77155	0.76733	0.8955	0.8955	0.8955	na
Gain in labor market earnings for former users vs current disordered users, probability density distribution parameters	Distribution Type	Gamma	Gamma	Gamma	Gamma	Gamma	Normal
	Alpha/Mean	39.608	53.157	47.014	47.014	47.014	1.9276
	Beta/Std Dev.	0.01255	0.00890	0.0052	0.0052	0.0052	0.03059
	Shift	0.77155	0.76733	0.8955	0.8955	0.8955	na

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:

$$(4.48)\ \ PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta ATOD_y \times \left(1 - \sum_{o=1}^{y} O_o\right) \times EarnGN \times EarnC_y\right) + \left(\Delta ATOD_y \times \left(1 - \left(1 - \sum_{o=1}^{y} O_o\right)\right) \times EarnGF \times EarnC_y\right)}{(1+dis)^{(y-tage+1)}}$$

Where $\Delta ATODy$ is the change in ATOD probability; 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 the following equation:

$$(4.49)\ \ PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta ATOD_y \times EarnGF \times EarnC_y\right)}{(1+dis)^{(y-tage+1)}}$$

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.

We also model the change in expected labor market earnings due to mortality. The present value of future labor market earnings at each age is multiplied by the decrease in probability that a person dies as the result of the disorder given that they have the disorder at that particular age.

4.4e Medical Costs, Treatment Costs, and Other Costs From ATOD

The WSIPP 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. Smoking health care costs are calculated with a different methodology explained later in this section.

Exhibit 4.53

Health Care Costs for ATOD Disorders

	Alcohol	Cannabis	Opioid drugs	Illicit drugs
Hospital-related parameters				
Annual number of disorder FTE hospital events (2012)[1]	18,905	n/a	4,897	14,406
Average charge per disorder FTE event (2012)[2]	36,150	n/a	41,663	36,895
SD of charge per disorder FTE event	49,628	n/a	62,471	79,980
Emergency department-related parameters				
Proportion of admissions attributable to disorder (2011)	7.9%	0.4%	0.7%	1.1%
Average ED expenses per admission (2011)	985	985	985	985
SE of average ED expense per admission	28.3	28.3	28.3	28.3
Treatment parameters				
Annual number treated (2013)	15,046	8,978	11,684	9,206
Average Cost per treatment episode (2015 dollars)	2,156	2,074	3,620	2,413
SD of Average cost per treatment episode (2015 dollars)	2,295	2,917	4,617	3,336

[1] *FTEHospitalEvent*
[2] *HospCostEvent*

Hospital-Related Parameters. The costs of hospital charges attributable to alcohol or illicit drugs 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 2012 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 Section 4.4c of this Chapter, 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*, given a stay. These parameters are shown in Exhibit 4.53. We also apply a hospital cost-to-charge ratio as described in Chapter 4.9.

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 heavy drinkers for some alcohol-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 using the following equations:

$$(4.50)\quad ExpHospEventsUpperBound = \frac{FTEHospitalEvents}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y-1}^{100} Pop_y}}$$

$$(4.51)\quad ExpHospEventsLowerBound = \frac{FTEHospitalEvents}{CurrentUse\% \ \times \sum_{y=1}^{100} Pop_y}$$

$$(4.52)\quad ExpHosp\$ = \frac{ExpHospEventUpperBound + ExpHospEventLowerBound}{2} \times HospCostEvent \ \times CostRatio$$

In computations, the upper bounds and lower bounds are averaged 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., (2008), we also make an adjustment to include pharmacological drugs and other medical non-durable costs. 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 4.53.

Emergency Department Parameters. Emergency department parameters are shown in Exhibit 4.53 for alcohol and drugs. 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 or drug related event. McDonald et al. (2004) estimate 7.9% of emergency room visits are alcohol related and data from the Drug Abuse Warning Network provide a national estimate of illicit drug-related emergency department visits of 0.64%, cannabis-related ED visits of 0.36%, and heroin and other opioid drug-related ED visits of 0.71%[118]

The total number of emergency department visits in Washington during 2008 is entered in Exhibit 4.53. These data come from a report by the Washington State Hospital Association.[119] We then apply the fractions just described; for example, for DSM alcohol disorders, we apply the 7.9% 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 0bound). We then apply a cost per emergency department event, *EDCostEvent*, and an emergency department cost-to-charge ratio. The average and standard error of the cost per emergency department visit is taken from the Medical Expenditure Panel Survey (MEPS) of the US Department of Health & Human Services.[120] In computations, the upper bounds and lower bounds are averaged to attribute an emergency department charge to a disordered DSM ATOD event (or heavy drinking episode where applicable), as given by the following equation:

$$(4.53)\quad ExpEDEventsUpperBound = \frac{TotalEDVists \times CausationFraction}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y-1}^{100} Pop_y}}$$

$$(4.54)\quad ExpEDEventsLowerBound = \frac{TotalEDVists \times CausationFraction}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$(4.55)\quad ExpED\$ = \frac{ExpEDEventsUpperBound + ExpEDEventsLowerBound}{2} \times EDCostEvent \times CostRatio$$

Treatment Parameters. For the costs of admissions to treatment, WSIPP was supplied with numbers by the Washington Department of Social and Health Services (DSHS). Number of admissions comes from the Treatment and Assessment Report Generation Tool (TARGET) database for FY 2013.[121] The TARGET database tracks patient instances and services. DSHS applied the modern public cost per treatment rate for each admission's course of treatment type by county and provider to estimate an average and standard deviation for the cost of treatment by type of substance. We assume that those admitted for treatment are part of the current annual prevalence of DSM disorders in the general population. We use the following equation:

$$(4.56)\quad ExpTreatmentEvents = \frac{TotalTreatmentEvents}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y-1}^{100} Pop_y}}$$

$$(4.57)\quad ExpTreatment\$ = ExpTreatmentEvents \times TreatmentCostEvent$$

Traffic Crash Parameters. We model alcohol-involved property costs with a similar set of procedures. We estimate the annual number of alcohol involved traffic crashes in Washington by obtaining the total number of officer-reported traffic collisions in Washington in 2011 (98,820).[122] To estimate the proportion of all crashes that are reported by police out of

[118] McDonald et al., (2004).
[119] Washington State Hospital Association. (2010). *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 multiplying by 12/18.
[120] Agency for Healthcare Research and Quality. (2011). Total Utilization and Mean Expenses per Visit by Type of Ambulatory Health Care Service, 2011 (Medical Expenditure Panel Survey Household Component Data). From http://meps.ahrq.gov/mepsweb/data_stats/summ_tables/hc/mean_expend/2011/table1.pdf. Retrieved April 7, 2014.
[121] Information from the TARGET database was provided via personal communication with Kevin Campbell, DSHS, May 12, 2016.
[122] Washington State Department of Transportation. (n.d.). *2011 Washington State collision data summary*. Olympia, WA: Author, Table 8. Retrieved Feb 26, 2013 from http://www.wsdot.wa.gov/mapsdata/collision/pdf/Washington_State_Collision_Data_Summary_2011.pdf

total crashes, we use national estimates produced by Blincoe et al., (2002).[123] Data from Blincoe provide an estimate that 56.7% of all crashes are reported by police.[124] Thus, an estimate of total crashes in Washington in 2011 is 174,267. To this we apply the alcohol induced causation factor (8.5%) derived from national information also provided in Blincoe et al., (2002), along with the average traffic crash cost, also from Blincoe et al. (2002) of $1,892 in 2000 dollars (see Exhibit 4.54).

Exhibit 4.54

Calculation of Average Property Costs from Alcohol-Caused Traffic Collisions

Collision category	Unit price in 2000 dollars	Total alcohol caused incidence	Percent of all crashes caused by alcohol
Property damage only	1,484	1,963,718	0.083
MAIS 0	1,019	183,511	0.072
MAIS 1	3,844	254,989	0.055
MAIS 2	3,954	72,082	0.165
MAIS 3	6,799	25,763	0.205
MAIS 4	9,833	6,502	0.178
MAIS 5	9,446	3,047	0.322
Fatal	10,273	13,570	0.325
Average	**1,892**		**0.085**

Source: Tables 12 and 13 of Blincoe et al., (2002).

From these inputs, we then compute an upper bound number of events per alcohol disorder under the assumption that all alcohol traffic events stemmed from individuals currently diagnosed with a DSM alcohol disorder (or heavy drinkers). A lower bound is calculated assuming that all alcohol related traffic events stemmed from any use of ATOD, whether or not the use was by a person with a DSM alcohol disorder (or heavy drinker) population using the following equations:

$$(4.58)\quad ExpTrafficCollisionsUpperBound = \frac{TotalTrafficCollisions \times CausationFraction}{\dfrac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y-1}^{100} Pop_y}}$$

$$(4.59)\quad ExpTrafficCollisionsLowerBound = \frac{TotalTrafficCollisions \times CausationFraction}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$(4.60)\quad ExpTrafficCollison\$ = \frac{ExpTrafficCollisionsUpperBound + ExpTrafficCollisionsLowerBound}{2} \times TrafficCostEvent$$

Smoking Health Care Cost Parameters. Smoking attributable health care costs were estimated using a pooled dataset from the 2007-2010 National Health Interview Survey (NHIS) linked to 2008-2011 Medical Expenditure Panel Survey. As explained in more detail in Section 4.8, MEPS data include a representative sample of NHIS households with additional detail collected on individual healthcare utilization and medical expenditures. We follow methodology outlined by Xu, et al., (2015)[125] in constructing a two-part model that examines smoking-attributable healthcare spending controlling for sociodemographic characteristics and other health-related behaviors and attitudes.

Two separate models were included in this analysis—a *prevention* model that estimated costs for non-smokers[126] compared to adults with *any* history of smoking (current or previous), and a *treatment* model that examined costs for former smokers relative to current smokers. Both models adjusted for demographic factors (age, sex, race/ethnicity, marital status); income/education factors (high school/college completion, poverty status, insured); health indicators (self-

[123] Blincoe, L.J., Seay, A.G., Zaloshnja, E., Miller, T.R., Romano, E.O., Luchter, S., & Spicer, R.S. (2002). *The economic impact of motor vehicle crashes 2000*. Washington, DC: United States Department of Transportation, National Highway Traffic Safety Administration.
[124] Ibid, table 3.
[125] Xu, X., Bishop, E.E., Kennedy, S.M., Simpson, S.A., & Pechacek, T.F. (2015). Annual healthcare spending attributable to cigarette smoking: An update. *American Journal of Preventive Medicine, 48*(3), 326-333.
[126] Note: non-smokers are defined as individuals that smoked less than 100 cigarettes during lifetime.

reported body mass index—overweight/obese, alcohol consumption/excessive drinking); and health related behaviors or attitudes (obtained flu shot in last year, wear seatbelt regularly, propensity to take risks, belief in ability to overcome illness without medical help). Medical comorbidities are not included in the model since smoking can exacerbate a wide-range of health conditions and can lead to multiple diseases, including cancer, chronic obstructive pulmonary disease (COPD), cardiovascular disease and diabetes.[127]

The first part of the estimating equation includes a logit model that determines the likelihood of any smoking (prevention model) or remaining a smoker versus becoming a former smoker (treatment model). In the second part of the model, total healthcare expenditures are estimated conditioned on entering the specified smoking status. The dependent variable, total healthcare expenditures, included costs related to hospital inpatient care, hospital outpatient care, office-based medical provider services, emergency department services and prescriptions. All cost estimates were converted to 2011 dollars using the Consumer Price Index (CPI)—Medical Component. The prevention and treatment models are shown below in Exhibits 4.55 and 4.56.

After deriving adjusted values for the overall effect of smoking on healthcare expenditures using the marginal effects, we create age-based estimates for the differential cost impact of smoking from age 18 to age 85. Standard errors of the estimates at each age are calculated by resampling the marginal distribution at each age and calculating the average of the standard deviations of the distributions. Exhibit 4.55 shows the average annual cost and incremental cost by year for prevention and treatment populations.

Exhibit 4.55

Input Parameters for the Incremental Health Care Costs of Smoking

	Prevention	Treatment
Annual incremental cost of disorder	$1,449.49	$358.91
Standard error on annual cost	$235.59	$476.75
Year of dollars	2011	2011
Age at which cost was measured	53	55
Additional cost per year of life beyond measurement age	$21.68	$7.84
Standard error on additional cost	$1.64	$3.15

Exhibit 4.56

Two-Part Model Assessing Healthcare Costs of Current or Former Smokers Relative to Never Smokers

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of smoking				
Age		0.03	***	(0.02 - 0.03)
Female		1.01	***	(0.89 - 1.14)
Race/ethnicity (ref: Hispanic)	White, non-Hispanic	0.71	***	(0.57 - 0.85)
	Black, non-Hispanic	0.33	***	(0.16 - 0.49)
	Asian, non-Hispanic	-0.09		(-0.35 - 0.17)
	Other, non-Hispanic	0.57	*	(-0.04 - 1.18)
Education (ref: Less than HS)	High school	0.03		(-0.13 - 0.19)
	Some college/AA	0.29	***	(0.14 - 0.45)
	College graduate/BA or higher	0.56	***	(0.36 - 0.76)
Marital status (ref: Married)	Never married, not cohabitating	-0.09		(-0.24 - 0.05)
	Divorced, separated, widowed	-0.02		(-0.18 - 0.15)
Poverty level (ref: below poverty level)	Near poor (100% to LT 125%)	-0.29	**	(-0.56 - -0.03)
	Low income (125% to LT 200%)	-0.16	*	(-0.35 - 0.03)
	Middle income (200% to LT 400%)	-0.18	**	(-0.35 - -0.02)
	High income (GE 400%)	0.22	**	(0.04 - 0.41)
Drinking status (ref: non-drinker)	Non-excessive drinker	0.03		(-0.14 - 0.19)
	Excessive drinker	0.06		(-0.12 - 0.25)

[127] United States. (2012). *Preventing tobacco use among youth and young adults: A report of the Surgeon General*. Rockville, MD: US Dept. of Health and Human Services, Public Health Service, Office of the Surgeon General.

	Unknown	0.58	*	(-0.08 - 1.24)
BMI group (ref: underweight)	Normal weight	0.24		(-0.24 - 0.71)
	Overweight	0.27		(-0.22 - 0.76)
	Obese	0.46	*	(-0.04 - 0.97)
Insured		1.03	***	(0.9 - 1.16)
Flu shot		0.8	***	(0.64 - 0.96)
Wear seatbelt	Always, nearly always	0.07		(-0.57 - 0.72)
	Sometimes, seldom/never	0.08		(-0.6 - 0.75)
Propensity to take risks	Uncertain-strongly disagree	-0.48		(-1.19 - 0.22)
	Agree somewhat/strongly	-0.47		(-1.17 - 0.24)
Belief in ability to overcome disease without medication	Uncertain-strongly disagree	0.5		(-0.28 - 1.28)
	Agree somewhat/strongly	0.14		(-0.66 - 0.93)
Smoke history		**0.06**		**(-0.08 - 0.2)**
Intercept		-1.67	***	(-2.55 - -0.78)
Part two: GLM, estimated costs				
Age		0.01	***	(0.01 - 0.02)
Female		0.09	**	(0.01 - 0.18)
Race/ethnicity (ref: Hispanic)	White, non-Hispanic	0.13	*	(-0.01 - 0.26)
	Black, non-Hispanic	0.1		(-0.04 - 0.25)
	Asian, non-Hispanic	-0.26	***	(-0.45 - -0.07)
	Other, non-Hispanic	0.27		(-0.09 - 0.64)
Education (ref: Less than HS)	High school	0.1		(-0.03 - 0.23)
	Some college/AA	0.02		(-0.09 - 0.12)
	College graduate/BA or higher	0.08		(-0.06 - 0.22)
Marital status (ref: Married)	Never married, not cohabitating	0		(-0.1 - 0.09)
	Divorced, separated, widowed	0.09	**	(0 - 0.18)
Poverty level (ref: below poverty level)	Near poor (100% to LT 125%)	-0.11		(-0.29 - 0.06)
	Low income (125% to LT 200%)	-0.08		(-0.21 - 0.05)
	Middle income (200% to LT 400%)	-0.22	***	(-0.34 - -0.1)
	High income (GE 400%)	-0.2	***	(-0.34 - -0.05)
Drinking status (ref: non-drinker)	Non-excessive drinker	-0.14	***	(-0.24 - -0.05)
	Excessive drinker	-0.35	***	(-0.47 - -0.23)
	Unknown	-0.27		(-0.67 - 0.13)
BMI group (ref: underweight)	Normal weight	-0.17		(-0.48 - 0.15)
	Overweight	-0.07		(-0.38 - 0.24)
	Obese	0.14		(-0.16 - 0.44)
Insured		0.34	***	(0.19 - 0.48)
Flu shot		0.24	***	(0.14 - 0.34)
Wear seatbelt	Always, nearly always	-0.79	***	(-1.12 - -0.46)
	Sometimes, seldom/never	-0.8	***	(-1.16 - -0.44)
Propensity to take risks	Uncertain-strongly disagree	0.09		(-0.31 - 0.5)
	Agree somewhat/strongly	0.05		(-0.36 - 0.47)
Belief in ability to overcome disease without medication	Uncertain-strongly disagree	0.09		(-0.31 - 0.5)
	Agree somewhat/strongly	-0.38	*	(-0.79 - 0.02)
Smoke history		**0.25**	*****	**(0.17 - 0.32)**
Intercept		8.28	***	(7.79 - 8.78)

no. of obs = 17,899
weighted size = 513,466,894
Design df = 204
F(30, 175) = 55.98
Prob > F =0.0000

Exhibit 4.57

Two-Part Model Assessing Healthcare Costs of Current Smokers Relative to Former Smokers

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of remaining a smoker				
Age		0.03	***	(0.02 - 0.04)
Female		1.06	***	(0.84 - 1.27)
Race/ethnicity (ref: Hispanic)	White, non-Hispanic	0.64	***	(0.43 - 0.85)
	Black, non-Hispanic	0.28	*	(-0.01 - 0.57)
	Asian, non-Hispanic	-0.04		(-0.49 - 0.41)
	Other, non-Hispanic	0.84	*	(0 - 1.69)
Education (ref: Less than HS)	High school	0.04		(-0.18 - 0.27)
	Some college/AA	0.26	**	(0.03 - 0.49)
	College graduate/BA or higher	0.34	**	(0 - 0.68)
Marital status (ref: Married)	Never married, not cohabitating	-0.01		(-0.24 - 0.21)
	Divorced, separated, widowed	-0.02		(-0.26 - 0.23)
Poverty level (ref: below poverty level)	Near poor (100% to LT 125%)	-0.5	**	(-0.93 - -0.07)
	Low income (125% to LT 200%)	-0.24		(-0.54 - 0.05)
	Middle income (200% to LT 400%)	-0.28	**	(-0.54 - -0.01)
	High income (GE 400%)	-0.11		(-0.4 - 0.19)
Drinking status (ref: non-drinker)	Non-excessive drinker	-0.16		(-0.46 - 0.13)
	Excessive drinker	0.05		(-0.22 - 0.31)
	Unknown	1.01	*	(-0.19 - 2.2)
BMI group (ref: underweight)	Normal weight	0.25		(-0.45 - 0.96)
	Overweight	0.35		(-0.37 - 1.08)
	Obese	0.57		(-0.15 - 1.29)
Insured		1.17	***	(0.97 - 1.37)
Flu shot		0.86	***	(0.57 - 1.15)
Wear seatbelt	Always, nearly always	-0.61		(-1.57 - 0.36)
	Sometimes, seldom/never	-0.52		(-1.51 - 0.48)
Propensity to take risks	Uncertain-strongly disagree	-1.16	*	(-2.51 - 0.19)
	Agree somewhat/strongly	-1.12		(-2.5 - 0.27)
Belief in ability to overcome disease without medication	Uncertain-strongly disagree	0.81		(-0.55 - 2.17)
	Agree somewhat/strongly	0.43		(-0.92 - 1.79)
Smoke current		**-0.37**	***	**(-0.58 - -0.15)**
Intercept		-0.14		(-1.42 - 1.14)
Part two: GLM, estimated costs				
Age		0.01	***	(0.01 - 0.02)
Female		0.07		(-0.05 - 0.2)
Race/ethnicity (ref: Hispanic)	White, non-Hispanic	0.1		(-0.12 - 0.32)
	Black, non-Hispanic	0.06		(-0.17 - 0.29)
	Asian, non-Hispanic	-0.15		(-0.52 - 0.23)
	Other, non-Hispanic	0.15		(-0.29 - 0.6)
Education (ref: Less than HS)	High school	0.17	**	(0 - 0.33)
	Some college/AA	0.03		(-0.11 - 0.18)
	College graduate/BA or higher	0.08		(-0.11 - 0.28)
Marital status (ref: Married)	Never married, not cohabitating	-0.07		(-0.22 - 0.08)
	Divorced, separated, widowed	0.07		(-0.07 - 0.2)
Poverty level (ref: below poverty level)	Near poor (100% to LT 125%)	-0.18		(-0.42 - 0.06)
	Low income (125% to LT 200%)	-0.14		(-0.32 - 0.04)
	Middle income (200% to LT 400%)	-0.2	**	(-0.35 - -0.05)
	High income (GE 400%)	-0.21	**	(-0.42 - -0.01)
Drinking status (ref: non-drinker)	Non-excessive drinker	-0.16	**	(-0.3 - -0.01)
	Excessive drinker	-0.39	***	(-0.58 - -0.21)
	Unknown	-0.32		(-0.82 - 0.18)
BMI group (ref: underweight)	Normal weight	0.01		(-0.34 - 0.35)
	Overweight	0.16		(-0.21 - 0.54)
	Obese	0.34	*	(-0.01 - 0.68)

Insured		0.24	**	(0.02 - 0.46)
Flu shot		0.37	***	(0.23 - 0.5)
Wear seatbelt	Always, nearly always	-0.73	***	(-1.05 - -0.41)
	Sometimes, seldom/never	-0.64	***	(-1.03 - -0.25)
Propensity to take risks	Uncertain-strongly disagree	-0.22		(-0.8 - 0.36)
	Agree somewhat/strongly	-0.28		(-0.85 - 0.3)
Belief in ability to overcome disease without medication	Uncertain-strongly disagree	0.45		(-0.14 - 1.04)
	Agree somewhat/strongly	-0.03		(-0.61 - 0.55)
Smoke current		**0.08**		**(-0.06 - 0.21)**
Intercept		8.36	***	(7.73 - 9)

no. of obs = 18,789 [subpop 7,458]
weighted size = 552,685,474 [subpop 225,196,485]
Design df = 204
F(30, 175) = 28.11
Prob > F =0.0000

4.4f Early Initiation of ATOD

As described above, we estimate the costs of disordered use of alcohol, cannabis, opioids, 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 early use of ATOD (e.g., by the end of middle school or the end of high school). 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 use of ATOD at early ages and the ultimate disordered use of ATOD. To estimate the relationship between early use and later disordered use of alcohol, cannabis, illicit drugs, and tobacco (regular use is the outcome of interest in the last case), we review the literature and contribute original analysis using NESARC data. Our estimates and sources for these early initiation parameters are described in Exhibit 4.58.

Exhibit 4.58

	Alcohol (a)	Cannabis (b)	Illicit drugs (non cannabis) (c)	Regular tobacco smoking (d)
Early initiation parameters:				
Prevalence of substance use by middle school [(1)]	29.5%	15.2%	8.7%	15.5%
Prevalence of substance use by high school [(2)]	69.4%	45.2%	24.1%	39.5%
D-cox effect size (ES) between early initiation and later disorder[(3)]				
Substance use by middle school	0.5472	1.29	1.6955	0.9583
Substance use by high school	0.6994	1.4263	1.762	1.3117
Standard error on d-cox ES between early initiation and later disorder				
Substance use by middle school	0.0462	0.1025	0.1177	0.0362
Substance use by high school	0.018	0.0574	0.063	0.018

1) Johnston, L.D., O'Malley, P.M., Bachman, J.G., & Schulenberg, J.E. (2013). *Monitoring the Future national survey results on drug use, 1975-2012: Volume I, Secondary school students.* Ann Arbor: Institute for Social Research, The University of Michigan. Middle school estimate from 8th grade results, Table 4-1a.
2) *Ibid.*High school estimate from 12th grade results, Table 4-1a.
3) Analysis of Wave 1 NESARC data. We computed a logistic regression coefficient for each ATOD, restricting the dataset to those people who ever used that substance, for those who first used the substance by age 14 (for the middle school analysis) or by age 18 (for the high school analysis) versus all others who first used that substance at a later age. This analysis controlled for age, sex, race/ethnicity, antisocial behavior by age 15, depression by age 14, and use of other substances. We then exponentiated these coefficients to obtain an odds ratio for early use versus later use. We then adjusted those odds ratios to account for the fact that we had "ever" users in our analysis by dividing the original odds ratios by the proportion of people in the general population who "ever" used that particular substance. From the adjusted odds ratios, we computed the input effect sizes between early use and later disordered use for each substance, and used @Risk software to estimate standard errors around those effect sizes.

4.5 Valuation of Teen Birth Outcomes

In the WSIPP benefit-cost model, the implications of a teen birth are expressed in terms of the birth's effect on long-term outcomes for the mother and child. 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 estimate these effects for both teen mothers and the children born to them.[128] The results from our meta-analyses of the research literature are shown in the Appendix. Our teen birth base rate number comes from the Washington Department of Health Vital Statistics and Population Data.[129] Because the teen birth rate has been trending downward in recent years, we use the most recent data available (2014), which shows a rate of approximately 8.3 teen births per 1,000 women.

4.6 Valuation of Public Assistance Outcomes

A portion of 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 by a family, 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 differences in this transfer are the effect that a change in public assistance caseloads has on costs related to the administration of the public assistance programs and the deadweight cost of the government taxation necessary to fund the transfer and its associated administrative costs.

4.6a Cash Assistance

We include the federal Temporary Assistance for Needy Families (TANF) and the state-run State Family Assistance (SFA) programs in the estimates of our value of cash assistance. We estimate the additional costs of public assistance cash transfers on a per-participant basis. Using state data reported to the federal Administration on Children and Families, we compute the total non-cash-assistance TANF expenditures as a proportion of total assistance expenditures.[130] These non-assistance costs include the cost of administering the program, as well as the cost of other, non-cash services that benefit TANF recipients. We compute the ratio of the non-assistance expenditures to the cash benefit on a per-participant basis to create the "Administrative proportion" shown in Exhibit 4.59. To estimate the proportion of total TANF/SFA expenditures that come from state versus federal sources, we use data reported by the TANF program.

4.6b Food Assistance

To estimate the value of food assistance, we include data from the federal Supplemental Nutrition Assistance Program (SNAP) and the state-run Food Assistance Program (FAP). Most of the costs of these programs are treated as transfer payments, similar to cash assistance. As SNAP and FAP do not provide other, non-cash-assistance services, any additional costs of these programs are the costs to administer the program.

Exhibit 4.59 displays the inputs for this area. Program effects for both cash assistance and food assistance are measured, most often, as a continuous measure of the number of months receiving assistance. Therefore, in addition to additional program costs and the proportion of state and federal expenditures, we also enter information on Washington State public assistance caseloads including the mean number of months on cash and food assistance for those on the caseloads, 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.

---

[128] 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.

[129] Retrieved August, 2015 from DOH Age-specific Live Birth Rates by Place of Residence. http://www.doh.wa.gov/DataandStatisticalReports/VitalStatisticsandPopulationData/Birth/BirthTablesbyTopic. We use the birth numbers for those ages 15-17 from table A10.

[130] Retrieved October 15, 2015 from http://www.acf.hhs.gov/programs/ofa/resource/tanf-financial-data-fy-2014 . Advice on categories to exclude (expenditures that would not be expected to be reduced if the adult caseload reduced) was provided via personal communication with Steve Ebben, Economic Services Administration, August 28, 2015.

We model a change in the number of months as the standard deviation change in the number of months spent receiving public or food assistance for those who receive assistance. The increase in months receiving benefits is multiplied by the average amount of monthly benefits in base year dollars. We know that the increase in months spent receiving public assistance occurs between the age of treatment and the age of measurement, so the total increase in assistance is evenly divided among all years between the age of treatment and the age at first measurement.

Exhibit 4.59

Public Assistance Parameters

	Cash assistance	Food assistance
Average monthly benefit	$398.56[1]	$222.62[2]
Administrative proportion	1.35[3]	0.11[4]
Average months on assistance	13.4[5]	40.5[6]
SD of months on assistance	16.0[5]	36.8[6]
Age at which assistance begins	18	18
Year of dollars	2014	2014
Proportion from state sources	0.269[7]	0.054[8]
Proportion from local sources	0.000[7]	0.000[8]
Proportion from federal sources	0.731[7]	0.946[8]

[1] Total dollars for TANF/SFA Regular Adult Cases divided by total cases for FY2015. Source: DSHS-ESA/EMAPS Assignment #3618 Using the ACES Data Warehouse as of September 2015.

[2] Total dollars for Total SNAP/FAP Cases divided by total cases for FY2015. Source: DSHS-ESA/EMAPS Assignment #3618 Using the ACES Data Warehouse as of September 2015.

[3] Total non-assistance TANF expenditures (net of the categories of "child care", "prevention of out of wedlock pregnancies," and "non-recurrent short-term benefits") divided by total assistance expenditures. Source: TANF Financial Data for FY2014 (http://www.acf.hhs.gov/programs/ofa/resource/tanf-financial-data-fy-2014). Advice on categories to exclude (expenditures that would not be expected to be reduced if the adult caseload reduced) was provided via personal communication with Steve Ebben, Economic Services Administration, August 28, 2015

[4] Monthly administrative costs divided by monthly household benefit, as reported in the SNAP State Activity Report, Fiscal Year 2014.

Source: http://www.fns.usda.gov/sites/default/files/FY14%20State%20Activity%20Report.pdf

[5] Total length of assistance and standard deviation in months computed using a cohort of adult clients entering TANF/SFA in January 2005 for the first time in Washington State. Source: ESA-EMAPS Report #3618 using the ACES Data Warehouse as of September 2015.

[6] Total length of assistance and standard deviation in months computed using a cohort of adult clients entering SNAP/FAP in January 2005 for the first time in Washington State. Source: ESA-EMAPS Report #3618 using the ACES Data Warehouse as of November 2015

[7] Proportion of costs borne by state and federal sources are derived from assistance and non-assistance categories reported in TANF Financial Data for FY2014 (http://www.acf.hhs.gov/programs/ofa/resource/tanf-financial-data-fy-2014), excluding the same categories as reported in note 3 above.

[8] Proportion of costs borne by state and federal sources are a weighted average of the breakdown of 1) administrative costs reported in the SNAP State Activity Report, Fiscal Year 2014 (http://www.fns.usda.gov/sites/default/files/FY14%20State%20Activity%20Report.pdf), and 2) direct benefit-costs reported by the Washington State Economic Services Administration (Source: DSHS-ESA/EMAPS Assignment #3618 Using the ACES Data Warehouse as of September 2015

4.7 Valuation of K–12 Education Outcomes

In valuing most K–12 education outcomes (i.e., standardized test scores, high school graduation, and years of education), we use a human capital approach, as described in Chapter 4.1. This section describes the inputs (Section 4.7a) and computational procedures (the subsequent subsections) we use to monetize those outcomes, as well as the methods for valuing two other outcomes of K–12 education frequently measured in the program evaluation literature: the use of special education and grade retention.

4.7a Education Parameters

Evaluations of education and other programs or policies often assess outcome measures such as student test scores, years of education, graduation rates, special education, or grade retention. WSIPP's benefit-cost model includes a number of education-related parameters used to compute estimates of the benefits of these education outcomes. The inputs entered into the model are shown in Exhibit 4.60. This subsection lists the individual inputs and their data sources.

Exhibit 4.60
General K–12 Education Parameters

		All students	Low-income students
State high school graduation rate		0.781	0.68
Standard deviation for number completed years of education		2.40	2.40
Cost of a year of education (2014 dollars) for a student in regular education		$8,695	$10,212
Cost of a year of education (2014 dollars) for a student in special education		$18,417	$19,934
Percent of students using special education		0.13	0.17
Average numbers of years in special education, for those who receive it		4	4
Average age of first entry into special education		8	8
Percent of students retained for at least one year		0.098	0.163
Average number of years retained, for those retained		1	1
	Max	0.42	0.42
Multiplier for human capital economic externalities of education	Mode	0.37	0.37
	Min	0.125	0.125
Gain in earnings for a 1SD increase in test scores	Mean	0.095	0.095
	SE	0.031	0.031
Gain in earnings from an additional year of education	Mean	0.100	0.100
	SE	0.024	0.024
Gain in high school graduation probability from a 1 SD increase in test scores	Mean	0.079	0.117
	SE	0.001	0.002

The High School Graduation Rate. The model contains a user-supplied parameter of the high school graduation rate. WSIPP's entry is Washington State's most recently published "on-time" graduation rate as published by the Office of Superintendent of Public Instruction (OSPI).[131] The on-time rate is defined as the percentage of public school students who graduate from high school within four years. We record OSPI's rate for all students and for low-income students.[132] In addition, WSIPP uses a lower predicted high school graduation rate for the juvenile offender population.[133] When the benefit-cost model is run, the baseline high school graduation rate is used in conjunction with effect sizes from programs that measure changes in the dichotomously measured high school graduation rate.

[131] Office of Superintendent of Public Instruction. (2015). *Graduation and Dropout Statistics Annual Report: Appendix A*. Olympia, WA: Author. Retrieved April 20, 2016 from http://www.k12.wa.us/dataadmin/

[132] 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% of federal poverty guidelines are eligible for free meals. Students in households up to 185% of federal poverty guidelines are eligible for reduced-price meals. For more information visit: http://www.k12.wa.us/ChildNutrition/Programs/NSLBP/default.aspx

[133] The high school graduation rate for juvenile offenders is calculated as the simple average of a lower and upper bound. For the lower bound, we use a number reported by the Department of Social and Health Services in 2012; they estimate that 9% of students served by the Juvenile Rehabilitation in 9th grade in the 2005/2006 school year graduated from high school on time (Coker et al. (2012). *High School Outcomes for DSHS-Served Youth*. Olympia, WA. Retrieved April 15, 2016 from https://www.dshs.wa.gov/sites/default/files/SESA/rda/documents/research-11-181.pdf). For the upper bound, we use a number from a 2014 report by the United States Office of Juvenile Justice and Delinquency Prevention that used American Community Survey data to calculate a status drop-out rate of 40% for institutionalized 16-to-24 year-olds (suggesting a graduation rate of 60%); Sickmund, Melissa, and Puzzanchera, Charles (eds.). 2014. Juvenile Offenders and Victims: 2014 National Report. Pittsburgh, PA: National Center for Juvenile Justice. Retrieved April 15, 2016 from http://www.ojjdp.gov/ojstatbb/nr2014/downloads/chapter1.pdf.

The Standard Deviation in the Number of Completed Years of K–20 Education. We use microdata from the March 2009 Current Population Survey to calculate the standard deviation in the number of years of education attained by adults age 25 or older in the US who completed at least 7th grade. When the benefit-cost model is run, the standard deviation in the number of years of education is used in conjunction with effect sizes from programs that measure the change in the number of years of education.

Costs of Regular K–12 Education. The model requires an estimate of the marginal cost of a year of K–12 education and the year in which these dollars are denominated.[134]

Special Education Parameters. The model can also calculate the value of two other K–12 educational outcomes: years of special education and grade retention. For special education, the information is entered for the cost of a year of special education and the year in which the special education costs per year are denominated.[135] The model also contains a user-supplied parameter of the percentage of students in special education. WSIPP's entry is the percentage of Washington State students in special education in 2014-15 (13.4%).[136] 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.[137] to adjust Washington's special education rate to 16.5% for low-income students.[138] We also estimate the 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.[139].

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. WSIPP's entry is based on 2009 national rates (9.8% of all students and 16.5% of low-income students) calculated by the US Department of Education.[140] These rates have dropped in recent years; in 1995, 16% of US students had been retained in a grade level.[141]

Multiplier for Human Capital Economic Externalities of Education. The model contains minimum, modal, and maximum estimates measuring the external economic benefits of education. These values are shown in Exhibit 4.60. There is a fairly large economic literature on this topic, summarized in a chapter by McMahon in Brewer.[142] Analysts have studied the degree to which growth in the private returns to human capital produce spillover economic gains to the rest of an economy. The low value we use is the estimate contained in Acemoglu & Angrist (2000).[143] The modal value is the

---

[134] The cost of regular education estimate is from: Office of Superintendent of Public Instruction. (2016). *Financial reporting summary: Washington State School Districts and Educational Service Districts* (Fiscal Year September 1, 2014–August 31, 2015). Olympia, WA: Author, Table 4. Retrieved May 26, 2016 from http://www.k12.wa.us/safs/PUB/FIN/1415/Full%202014-15%20Financial%20Reporting%20Summary.pdf

[135] 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 special education allocation estimate is from: Office of Superintendent of Public Instruction. (2016). *Financial reporting summary: Washington State School Districts and Educational Service Districts* (Fiscal Year September 1, 2014–August 31, 2015).

[136] Office of Superintendent of Public Instruction. *Washington State Report Card*. Retrieved May 26, 2016 from http://reportcard.ospi.k12.wa.us/summary.aspx?groupLevel=District&schoolId=1&reportLevel=State&year=2014-15

[137] 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., et. al. (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

[138] We took the percentage of children in special education for up to 185% 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.

[139] The average number of years of special education and the average age of first entry in special education are WSIPP estimates.

[140] Planty et al. (2009), Table A-18-1, retrieved from: http://nces.ed.gov/pubs2009/2009081.pdf. For low income students, we computed a weighted average of poor (22.9%) and near-poor (10.9%) grade retention rates from table A-18-1. The weighting comes from the number of individuals in the respective groups taken from Table A-6-2.

[141] National Center for Education Statistics. (2006). *The Condition of Education 2006* (NCES 2006-071). Washington, DC: Author. Retrieved June 17, 2016 from http://nces.ed.gov/pubs2006/2006071.pdf.

[142] McMahon, M. (2010). The external benefits of education. In D.J. Brewer, & P.J. McEwan (Eds.) *Economics of education*. Oxford, UK: Academic Press.

[143] Acemoglu, D., & Angrist, J. (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual,* 15, 9-59.

estimate used in Belfield, Hollands, and Levin (2011).[144] The high parameter is contained in Bretton (2010).[145] In the model, a Monte Carlo draw is taken from a triangular probability density distribution with these three bounding parameters. The parameter is expressed as a multiple of the private economic return to education. For example, if the private return for a year of education is 0.10 and a modal external economic return parameter is 0.37, then the model monetizes the external economic benefits as 0.10 X 0.37 = 0.037 and this value is, in turn, multiplied times the valuation of the education-attributed difference in private earnings.

Fiscal Sources for Regular and Special Education Expenditures. As noted, the model allows users to input the proportion of education funding from state, local, and federal sources. While the model allows the user to enter separate values for the fiscal sources for regular- and low-income students, for Washington we enter the same figures for both. Washington State sources are described in Exhibit 4.61.

Exhibit 4.61

Proportion of Marginal Education Costs by Source

	State	Local	Federal
Regular education[1]	0.694	0.227	0.079
Special education[2]	0.851	0.000	0.149

[1] Washington State Office of the Superintendent of Public Instruction, "Statewide Average Financial Tables and Charts" for school year 2014-2015, Table 3, available at: http://www.k12.wa.us/safs/PUB/FIN/1415/Full%202014-15%20Financial%20Reporting%20Summary.pdf
[2] Washington State Office of the Superintendent of Public Instruction, "Statewide Average Financial Tables and Charts" for school year 2014-2015, General Fund Expenditures by Program, available at: http://www.k12.wa.us/safs/PUB/FIN/1415/Full%202014-15%20Financial%20Reporting%20Summary.pdf

4.7b Linkages: Education

WSIPP's benefit-cost model monetizes improvements in educational outcomes, in part, with linkages between each educational 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 high school graduation and crime 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 WSIPP model are listed in the Appendix. In addition, several relationships are modeled using the methods described below.

The Relationship Between Gains in Test Scores and the High School Graduation Rate. In many outcome evaluations of education programs, the only measure of effectiveness is student performance on standardized tests. In the WSIPP benefit-cost approach, however, we also model the likelihood of high school graduation where possible. Using Washington State data, we were able to estimate the increased likelihood of high school graduation, given improvement in standardized test scores. This additional analysis allows us to predict the impact of a program on high school graduation when evaluations of that program have only measured standardized test score performance. High school graduation, of course, is a marker for other student skills than just test scores, but performance on test scores is correlated with graduation.

We estimate the relationship between standardized test scores and high school graduation using longitudinal, student-level assessment and enrollment data for Washington State. These data include math and reading Washington Assessment of Student Learning (WASL) scores (in 7th, 8th, and 10th grades) for two cohorts of students (enrolled in 7th grade during 2004-05 or 2005-06). These students were expected to graduate in 2010 or 2011.

Three sets of models were run to examine the effects of: 1) changes in test scores between 7th and 8th grade; 2) changes in scores between 8th and 10th grades; and 3) test retake scores in 11th grade.[146] These models produced roughly

[144] Belfield, C., Hollands, F., & Levin, H. (2011). *What are the social and economic returns?* New York: Columbia University, Teachers College, The Campaign for Educational Equity.
[145] Breton, T.R. (2010). Schooling and national income: How large are the externalities? Corrected estimates. *Education Economics, 18*(4), 455-456.
[146] Many, but not all, students who did not meet assessment standards in 10th grade retake exams in 11th grade.

comparable estimates for the effect of assessment scores on graduation. The models that focus on 8th and 10th grade scores have the most observations, and we used these results for inputs to the benefit-cost model.

We ran linear probability models to estimate the effect of 10th grade test scores on graduation status, controlling for 8th grade test scores and other observed student characteristics.[147] The models did not fully control for unobserved student characteristics, and the extent to which estimates reflect cause-and-effect remains, to a degree, uncertain. For the analysis, the assessment scores were converted to Z-scores (mean 0, standard deviation 1). The difference in Z-scores between 8th and 10th grade reflects the change in a student's assessment scores. We estimated separate models for math and reading test scores. We also estimated separate models for low-income students.[148] Math estimates were based on observations for 114,221 students; reading estimates were based on data for 115,557 students. The basic equation estimated is shown below.

$$(4.61) \quad Graduation_i = \alpha + \beta_1 \Delta Z_i + \beta_2 \Delta Z_i \cdot Z_{8i} + \beta_3 Z_{8i} + \delta' X_i + \xi Year_i + \epsilon_i$$

Where:

$Graduation_i$ = 1 if student graduates, 0 if not
ΔZ_i = change in Z scores for student i = Z_{10i}-Z_{8i}
Z_{10i} = math (or reading) Z-score for 10th grade for student i
Z_{8i} = math (or reading) Z-score for 8th grade for student i
X_i = a vector of student characteristics (free or reduced-price meal eligibility history, English language status, special education status, gender, race/ethnicity)
$Year_i$ = indicator for the 10th grade assessment year

Exhibits 4.62 and 4.63 summarize the estimated effects of math and reading test scores on graduation status. The effects are determined by β_1 and β_2.[149] β_1 is the coefficient for the change in Z-scores. β_2 is the coefficient for an interaction term which allows the effect of test score growth to vary with the initial (8th grade) score.

Exhibit 4.62

Estimated Effects of Changes in Test Scores on Likelihood of High School Graduation, for All Students

	Math		Reading	
	Coefficient	Standard error	Coefficient	Standard error
ΔZ_i	0.0961	0.0021	0.0612	0.0015
$\Delta Z_i \cdot Z_{8i}$	-0.0172	0.0017	0.0001	0.0010

Note: The regression models also control for student characteristics and initial year test scores. Robust (to heteroskedasticity) standard errors are estimated.

Exhibit 4.63

Estimated Effects of Changes in Test Scores on Likelihood of High School Graduation, for Low-Income Students

	Math		Reading	
	Coefficient	Standard error	Coefficient	Standard error
ΔZ_i	0.1337	0.0033	0.0973	0.0026
$\Delta Z_i \cdot Z_{8i}$	-0.0046	0.0031	-0.0022	0.0017

Note: The regression models also control for student characteristics and initial year test scores. Robust (to heteroskedasticity) standard errors are estimated.

These regression results for math and reading were then averaged to provide the "test score" effect for the benefit-cost model, and these averages are entered in the model. The standard errors for the test score averages were calculated by running 10,000 case Monte Carlo simulations with the test score specific parameters in Exhibits 4.62 and 4.63.

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

[147] We estimate robust standard errors for the linear probability models. We also estimated logistic regression models and inferences were comparable.
[148] Low-income students are defined as ever having been eligible for free or reduced-price meals.
[149] The effect of a change in test score is given by $d(graduation)/d(\Delta Z) = \beta 1 + \beta 2 \cdot Z8i$.

for this input is used in Monte Carlo simulations (see Chapter 6). For these two parameters, we use regression results from Hall & Farkas (2011).[150] They estimate multi-level models of cognitive ability (measured with standardized test scores) and attitudinal/behavioral traits (sometimes called non-cognitive skills) on log wages with data from the National Longitudinal Survey of Youth (NLSY79). We compute weighted averages from their results for males and females, and for white, black, and Latino populations. We use Monte Carlo simulation to estimate a standard error from their constant and slope parameters. Their results are useful for the benefit-cost model because the cognitive ability scale they create measures several areas (word knowledge, paragraph comprehension, math knowledge, and arithmetic reasoning) often found in the program evaluation literature. The results from the Hall & Farkas study are in line, though slightly lower, than those found in other studies.[151] We enter the same parameter for all students and for low-income students because, to date, we have not found separate estimates for low-income populations. When additional research is conducted, separate estimates can be entered for low-income students.

The Relationship between Gains in Years of Secondary 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[%]."[152] 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.[153] 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% increase in labor market earnings per additional year of education completed (with a standard error of 0.02).[154] The study by Hall and Farkas (2011) that we use for the effect of student test scores on labor market earnings, found a 9.5% rate of return for an extra year of education—a rate very similar to the 10% rate we use in our model. 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. If and when additional research is conducted, separate estimates can be entered for low-income students.

The Relationship between High School Graduation and Labor Market Earnings. The model contains two types of parameters, both shown in Exhibit 4.64, to measure the labor market earnings effect of graduating from high school. The two types of parameters model the analytical framework established in a recent paper by Heckman et al.[155] One type of parameter is a high school graduation causal factor, which measures the degree to which the observed difference in earnings between types of high school graduates and non-high school graduates is causal and is derived from the Heckman 2014 analysis. These values and their standard errors are derived separately by the highest level of education completed. The second set of estimates measure the sequential probability that high school graduation opens the possibility of an individual continuing to obtain some additional college education or completes a college degree. These probabilities were calculated from the share of high school graduates with some college or a 4-year degree or higher as reported in the American Community Survey 2010-2014 for Washington State. The estimates represent the proportion of

[150] Hall, M. & Farkas, G. (2011). Adolescent cognitive skills attitudinal/behavioral traits and career wages. *Social Forces, 89*(4), 1261-1285.
[151] See 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.
[152] Psacharopoulos, G., & Patrinos, H.A. (2004). Returns to investment in education: A further update. *Education Economics, 12*(2), 111-134.
[153] Heckman, J., Lochner, P., & Todd, P. (2008). Earnings functions and rates of return. *Journal of Human Capital, 2*(1), 1-31.
[154] 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* (pp. 1801-1863). Atlanta, GA: Elsevier.
[155] Heckman et al., (2015).

those in Washington aged 25 and older with some college (no degree or any degree less than a Bachelor's) and those with a BA or greater. Unlike our previous estimates, we were unable to separate on-time high school graduates from those with late completions or GED attainment. We further assume that some high school certification is necessary to continue on to further levels of education.

Those who continue on to college incur the cost of college education. High school graduation is a pathway to further education and the associated costs. WSIPP estimates these costs per year of education, then multiplies these numbers by the average number of years that students spend in school to produce the stream of higher education costs for the some college and college graduate paths. We describe the calculation in detail in section 4.7c.

4.7c Valuation of Earnings from High School Graduation

The full equation for the value of a high school education is displayed in Equation 4.62.

Exhibit 4.64

Estimates of the Causal Effect of High School Graduation on Earnings

		High school graduate (only)	Some college	College graduate
Percent of high school graduates who go on to each level of education		0.26	0.38	0.36
Percent of observed earnings gains caused by high school graduation	Mean	0.50	0.56	0.42
	SE	0.17	0.13	0.11

$$
\begin{aligned}
(4.62)\quad & EarnGainHSG_y \\
&= \Big(\big(EarnHSG_y \times \%HSG \times HSGCF \times (1+EscHSG)^{y-age} \times (FHSG \times (1+EscFHSG)^{y-age}) \\
&\times StateAdjHSG\big) + (EarnSomeCol_y \times \%SomeCol \times SomeColCF \times (1+EscSomeCol)^{y-age} \\
&\times (FSomeCol \times (1+EscFSomeCol)^{y-age}) \times StateAdjSomeCol - CostSomeCol) + (EarnColDeg_y \\
&\times \%ColDeg \times EarnColDegCF \times (1+EscColDeg)^{y-age} \times (FColDeg \times (1+EscFColDeg)^{y-age}) \\
&\times StateAdjColDeg - CostColDeg) - (EarnNHSG_y \times (1+EscNHSG)^{y-age} \\
&\times (FNHSG \times (1+EscFNHSG)^{y-age}) \times StateAdjNHSG) \Big) \times (IPD_{base}/IPD_{cps}) \times (1+HCEXT)
\end{aligned}
$$

For each year (y) over the course of a person's working career, the expected earnings gain from graduating from high school versus not graduating from high school, *EarnGainHSG*, is the product of:

a) the observed earnings of high school graduates in each year, $EarnHSG_y$, times the percent of high school graduates who do not pursue further education, *%HSG*, times the high school graduation causation factor, *HSGCF*, times one plus the relevant real earnings escalation rate for high school graduates (*EscHSG*) raised to the number of years after program participation, times the fringe benefit rate for high school graduates (*FHSG*), times one plus the relevant fringe benefit escalation rate for all people (*EscFHSG*) raised to the number of years after program participation, times the ratio of state-to-national earnings for high school graduates (*StateAdjHSG*); plus

b) the observed earnings of people with some college in each year, $EarnSomeCol_y$, times the percent of high school graduates who pursue some college, *%SomeCol*, times the some college graduation causation factor, *SomeColCF*, times one plus the real earnings escalation rate for those who pursue some college (*EscSomeCol*) raised to the number of years after program participation, times the fringe benefit rate for those who pursue some college (*FSomeCol*), times one plus the relevant fringe benefit escalation rate for those who pursue some college (*EscFSomeCol*) raised to the number of years after program participation, times the ratio of state-to-national earnings for those with some college (*StateAdjSomeCol)* minus the cost of some college (*CostSomeCol);* plus

c) the observed earnings of people with college degrees in each year, $EarnColDeg_y$, times the percent of high school graduates who obtain a college degree, *%ColDeg*, times the college degree causation factor, *EarnColDegCF*, times one plus the real earnings escalation rate for those who obtain a college degree (*EscColDeg*) raised to the number of years after program participation, times the fringe benefit rate for those who obtain a college degree (*FColDeg*), times one plus the relevant fringe benefit escalation rate for those who obtain a college degree (*EscFColDeg*) raised to the number of years after program participation, times the ratio of state-to-national earnings for those with college degrees (*StateAdjColDeg)* minus the cost of a year of college for a college graduate (*CostColDegree);* minus

d) the observed earnings of people who do not graduate from high school in each year, $EarnNHSG_y$, times one plus the real earnings escalation rate of people who do not graduate from high school (*EscNHSG*) raised to the number of years after program participation, times the fringe benefit rate of people who do not graduate from high school (*FNHSG*), times one plus the relevant fringe benefit escalation rate of people who do not graduate from high school (*EscFNHSG*) raised to the number of years after program participation, times the ratio of state-to-national earnings for non-high school graduates (*StateAdjNHSG);*

e) the product is the multiplied by 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}., times one plus the parameter for economic gain from human capital externalities, *HCEXT.* [156]

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

$$(4.63)\ \ PVEarnGainHSG = \sum_{y=age}^{65} \frac{EarnGainHSG_y \times Units_{hsg}}{(1 + Dis)^{y-age}}$$

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" at age 18 (in percentage points), $Units_{hsg}$, caused by the program or policy. The calculation of the units variable is described in Chapter 2 and 3. 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.

Estimating costs of higher education and sources of revenues

Part of the benefit of the labor market gains from high school graduation comes from a college education. We estimate the costs of obtaining that education. We analyzed data from the Integrated Postsecondary Education Data System (IPEDS). The cost per year of higher education is estimated as the institutional expenditures per full-time equivalent (FTE) undergraduate student required to finance a student's education at each institution in Washington. The estimated cost per FTE includes expenditures for instruction, academic support, student services, institutional support, and operation and maintenance of plant (i.e. the physical institution).[157] Exhibit 4.65 shows our estimates for cost and payer by type of student and education.

To calculate the cost per undergraduate FTE in Washington, we weight graduate FTEs by an additional 25% as graduate students incur more costs than undergraduate students.[158] We then sum the included expenses for each of the 2- and 4-year institutions in Washington State and divide the sum by the total number of FTEs (with graduate students weighted more) to arrive at an average cost per undergraduate FTE for each institution. We average the costs per FTE across all institutions weighted by the number of undergraduates. We calculate this average for 2-year and 4-year institutions separately and overall. The estimate only using 4-year institutions are reported as "undergraduate graduates," while those using information from both 2-year and 4-year institutions are reported as the costs for "some college, no BA."

[156] During years when students are in college, we do not apply the externality multiplier to their decreased earnings relative to non-college attendees. That is, we do not monetize negative human capital externalities.
[157] We exclude expenses for research, public service, auxiliary, hospital services, independent operations, and other expenses. We also exclude scholarship and fellowship expenses that are paid for goods and services not provided by the institution (e.g. scholarships and fellowship expenses for off-campus housing).
[158] National Association of College and University Business Officers. (2002). Explaining College Costs: NACUBO's Methodology for identifying the costs of delivering undergraduate education.

Exhibit 4.65

Higher Education Parameters

	Some college, no BA		Undergraduate graduates	
	All students	Low-income students	All students	Low-income students
Annual cost, some college, no BA	$16,312	$16,312	$22,961	$22,961
SD cost	$8,501	$8,501	$9,414	$9,414
Year dollars	2014	2014	2014	2014
Number of years	2.48	2.48	4.07	4.07
Percent paid by participant	47%	37%	55%	43%
Percent paid by taxpayer	40%	53%	28%	41%
Federal	27%	21%	27%	23%
State	72%	79%	73%	77%
Local	0%	0%	0%	0%
Percent paid by others	14%	14%	18%	16%

To determine the share of revenues paid by students, taxpayers, and others, we first estimate revenues per FTE including only those revenues coming from state, federal, and local appropriations and grants given directly to students as scholarships or fellowships (e.g. Pell grants), institutional and private grants, and tuition revenue from students.[159] We divide these revenues by the number of FTEs to arrive at total funding per FTE.[160] We use the same methodology to calculate revenues per FTE coming from each source (i.e. state, federal, local, institutional/private, and students). We then divide funds from state, federal, local, other sources and from tuition revenue per FTE by the total amount of funding per FTE to estimate the share of total funds for education that are paid by each source.

The above methodology will provide an estimate of the share of revenues derived from each source for the average student. However, low-income students receive the bulk of state and federal grant funding as Pell Grants and Washington's State Need Grant are only available to low-income students. To estimate share of revenues from each source for low-income students, we use the IPEDS data on the financial aid cohort. IPEDS financial aid data provides information on the total amount of grant funding by income categories and the number of undergraduate students in each income category.[161] We use this information to approximate the average grant amount per FTE for those in lower income categories.

Because we do not have more granular income data for students, we define low income students as those with family incomes less than $48,000.[162] To estimate the total amount of revenues from each source going to low-income students, we multiply the total amount of state, federal, local grants, or institutional/private funds for all students by the percent of all grants and scholarship dollars going to low-income students.[163] We then divide this estimate of total grant funding to low-income students by source by the percent of undergraduates that are low income to arrive at the per low-income FTE amount of grant funding from state, federal, local, and institutional/private sources. The additional funding from these sources for low-income students is then subtracted from the tuition revenue to account for the fact that increased grant funding reduces the share students pay themselves.

---

[159] Contracts and grants for research are excluded from the grant funds as are other non-operating grants that are not provided to students to finance their educations. We also exclude revenues from auxiliary enterprises, independent operations, investment income, capital appropriations and grants, and private gifts.

[160] We divide the total amount of state and federal grant funding by the number of undergraduate FTEs, as this funding generally applies only to undergraduates.

[161] Note that because students from high income families may not apply for financial aid, using information from the financial aid cohort will probably overestimate the proportion of students that are low income.

[162] IPEDS income categories are $0-30K, $30-48K, $48-75K, $75-100K, and greater than $100,000. In Washington State, students at or below 70% of median income ($58,500) can receive State Need Grant funding. For Pell grants, students with family incomes below $50,000 can receive funding

[163] The IPEDS financial aid data does not provide information on the total amount of grant and scholarship funding broken out by source and income category. Data on total amount of grant and scholarship funding by income category is for all sources combined.

Higher education costs are increased by a long-run real escalation rate in per capita inflation-adjusted higher education costs.[164] The model uses a triangular distribution around three different estimates of real higher education costs escalation (low =0.000, modal =0.0083, high =0.0129). The escalation is applied beginning in the year following treatment.

For each year or partial year that a person spends in higher education, the expected cost of a year of college is the product of the percent of the year in school multiplied by the cost of that type of attendance (some college versus college graduate).

4.7d 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 via a "linked" outcome), we use the Current Population Survey (CPS) earnings data, described in Section 4.1, and the other parameters described in Section 4.7a, 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 CPS 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, as described in Chapter 4.1. 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}, times the ratio of state-to-national earnings for all people (*StateAdjAll*), times the degree of causation, *TSCF*, between a one standard deviation gain in student test scores and the related percentage increase in labor market earnings, times one plus the parameter for economic gain from human capital externalities, *HCEXT.*

$$(4.64)\ ModEarnAll_y = \left(EarnAll_y \times (1+EscAll)^{y-age}\right) \times \left(FAll \times (1+EscFAll)^{y-age}\right) \times \left(IPD_{base}/IPD_{cps}\right) \times StateAdjAll \times TSCF \times (1+HCEXT)$$

[164] The low estimate was based on the assumption that higher education costs grow at the same rate as other expenses. The middle and high estimates for escalation in the cost of higher education were computed by calculating the compound annual growth rate for the HECA and HEPI education cost indexes from 1989 to 2014 as reported in the technical paper of the State Higher Education Executive Officers Association. SHEEO. (2014). *The Higher Education Cost Adjustment: A Proposed Tool for Assessing Inflation in Higher Education Costs.* From http://www.sheeo.org/sites/default/files/SHEEO002_2014AdtlDocs_TechA_Rd1.pdf. Retrieved on April 6th, 2016. We use the differences between the CAGR of HECA/HEPI and the IPD (see section 4.10f) are used for the mid/high inputs.

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 test score units are measured at age 17. The calculation of the units variable is described in Chapters 2 and 3. 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, as given by the following equation:

$$(4.65)\ \ PVEarnGainTS = \sum_{y=age}^{65} \frac{ModEarnAll_y \times Units_{ts}}{(1+Dis)^{y-age}}$$

4.7e 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 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}, times the ratio of state-to-national earnings for all people (*StateAdjAll*), times the degree of causation, *YearsOfEdCF*, between one extra year of education related percentage increase in labor market earnings, times one plus the parameter for economic gain from human capital externalities, *HCEXT*

$$\begin{aligned}(4.66)\ \ & ModEarnAll_y \\ &= \left(EarnAll_y \times (1+EscAll)^{y-age}\right) \times \left(FAll \times (1+EscFAll)^{y-age}\right) \times \left(IPD_{base}/IPD_{cps}\right) \times StateAdjAll \\ &\quad \times YearOfEdCF \times (1+HCEXT)\end{aligned}$$

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 Chapters 2 and 3. 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.

$$(4.67)\ \ PVEarnGainYearsofEd = \sum_{y=age}^{65} \frac{ModEarnAll_y \times Units_{yearsofed}}{(1+Dis)^{y-age}}$$

4.7f 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, *IPDbase*, relative to the year in which the special education costs per year are denominated, *IPDspecedcostyear*.

$$(4.68)\ \ PVspeced_{start} = \sum_{y=1}^{specedyears} \frac{SpecEdCostYear}{(1+Dis)^{y}}$$

$$(4.69)\quad PVspeced_{progage} = \frac{PVspeced_{start} \times \dfrac{IPD_{base}}{IPD_{speced\,costyear}}}{(1+Dis)^{start-progage}}$$

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, $Hsgradprob$.

$$(4.70)\quad PVgraderet_{progage} = \left[\frac{EdCostYear \times \dfrac{IPD_{base}}{IPD_{ed\,costyear}}}{(1+Dis)^{18-progage}} \right] \times Hsgradprol$$

4.7g Adjustment 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 always remain constant over time; researchers have found that test score gains from program participation often get smaller (the test scores decay or "fade out") as years pass after the intervention.[165]

Most of the evaluations of educational interventions we examine in our meta-analyses measure test score performance in elementary school. However, the relationships in the economic literature between test scores and labor market earnings are based on test scores measured late in high school. Therefore, for use in the benefit-cost model, it is necessary to adjust earlier measurements of test scores appropriately in order to more accurately model the economic benefits resulting from improvements in standardized test scores measured in program evaluations. When we include test score effect sizes from evaluations of programs which measure scores in their pre-high school years, we apply a multiplicative adjustment to account for the average fadeout observed in research.

To estimate the magnitude of this fadeout for test scores measured at different points in time, we focus on research that follows children who attended state, district, home school, or model pre-kindergarten education programs and measured those children's scores on standardized tests for some period of time. The follow-up periods for test score measures in the 59 studies we analyzed varied widely. We conducted meta-analyses of effect sizes from these 59 studies covering four periods of time after the early childhood intervention: immediately after preschool, kindergarten–2nd grade, 3rd–5th grade, and 6th–9th grade (Exhibit 4.66). We included both IQ tests and standardized academic tests from specific program evaluations and national surveys.

[165] For example, a meta-analysis by Leak et al. (2010) found that early test score gains decreased by at least 54% 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% 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% per follow-up period. Leak, J., Duncan, G., Li, W., Magnuson, K., Schindler, H., & Yoshikawa H. (2010). *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; and Goodman, A. & Sianesi, B. (2005). Early education and children's outcomes: How long do the impacts last? *Fiscal Studies, 26*(4), 513-548.

Exhibit 4.66

Meta-Analytic Results at Four Time Periods

Time of measurement	Number of effect sizes	Average time since the beginning of preschool (years)	Average effect size	Standard error
Immediately after preschool	37	1	0.309	0.030
Kindergarten–2nd grade	38	2.9	0.152	0.019
3rd–5th grade	29	5.7	0.097	0.014
6th–9th grade	12	9.4	0.085	0.033

As seen in Exhibit 4.66, the average effect size measured immediately after preschool reduces significantly over time. The meta-analytic results suggest a non-linear relationship between the effect size and the time since the intervention. We tested the following models to fit a trend line to the data: quadratic, cubic, logarithmic, and power. A power curve provided the best combination fit (R^2=0.98) and a believable pattern of decay (Exhibit 4.67). The decrease in effect size by 3rd–5th grade was similar to that found by Camilli et al. (2010). We used the power curve model to estimate the effect sizes through 12th grade. We also modeled the relationship between the effect size and the time since the intervention using meta-regression. However, various model specifications led to notably different intercepts, thus we opted to use the simpler meta-analytic results to model fadeout. We projected these findings out to 12th grade for use in the benefit-cost model. Exhibit 4.68 displays the adjustment factors we use in the benefit-cost model.

Exhibit 4.67

Estimation of Test Score Fadeout:
Meta-Analytic Results and Power Curve Model

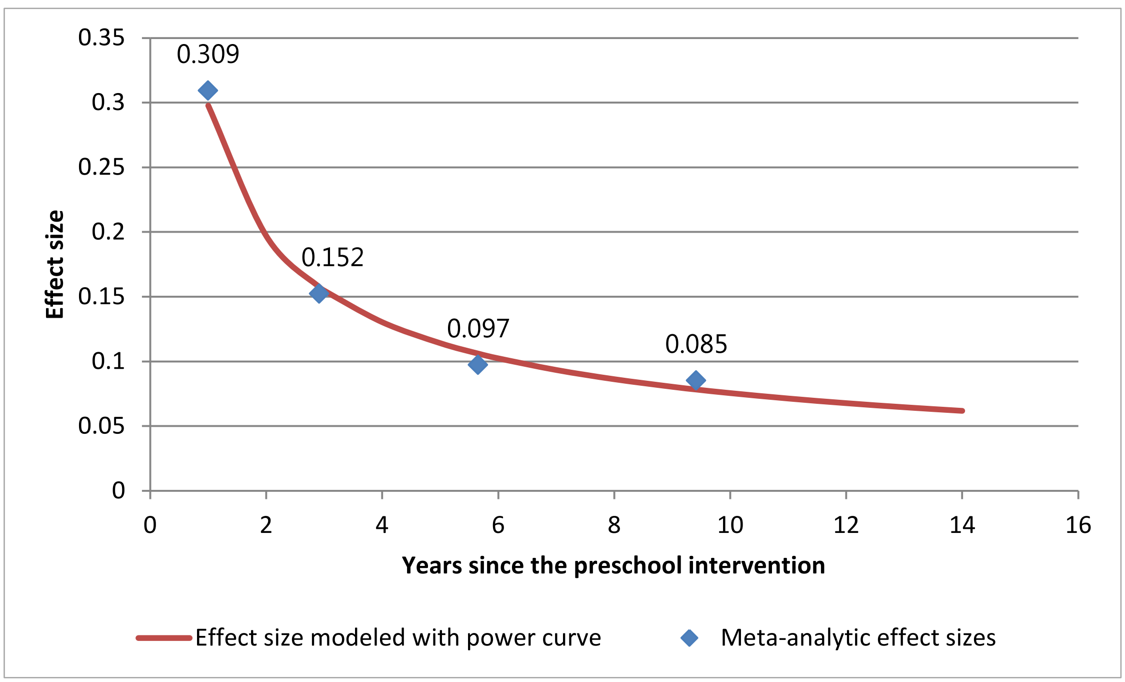


Exhibit 4.68

Fadeout Multipliers for Test Scores:
Estimates of Effect Size Decay Based on Longitudinal Evaluations of Early Childhood Education

Age at measurement	Grade level	Fadeout: Later test score effect size as a % of pre-K effect size	Fadeout multiplier: Multiply the effect size by the % below to estimate end-of-high school effect
4	Pre-K	100%	21%
5	K	66%	31%
6	1	52%	40%
7	2	44%	47%
8	3	38%	54%
9	4	34%	60%
10	5	31%	66%
11	6	29%	72%
12	7	27%	77%
13	8	25%	82%
14	9	24%	87%
15	10	23%	91%
16	11	22%	96%
17	12	21%	100%

Studies Used in Test Score Fadeout Analysis:

Abbott-Shim, M., Lambert, R., & McCarty, F. (2003). A comparison of school readiness outcomes for children randomly assigned to a Head Start program and the program's wait list. *Journal of Education for Students Placed at Risk, 8*(2), 191-214.

Andrews, R.J., Jargowsky, P.A., & Kuhne, K. (2012). *The effects of Texas's targeted pre-kindergarten program on academic performance.* Cambridge, MA: National Bureau of Economic Research.

Apps, P., Mendolia, S., & Walker, I. (2013). The impact of pre-school on adolescents' outcomes: Evidence from a recent English cohort. *Economics of Education Review, 37*, 183-199.

Aughinbaugh, A. (2001). Does Head Start yield long-term benefits? *The Journal of Human Resources, 36*(4), 641-665.

Barnett, W.S., Frede, E.C., Mobasher, H., & Mohr, P. (1988). The efficacy of public preschool programs and the relationship of program quality to efficacy. *Educational Evaluation and Policy Analysis, 10*(1), 37–49.

Barnett, W.S., Jung, K., Youn, M., & Frede, E.C. (2013). *Abbott preschool program longitudinal effects study: Fifth grade follow-up.* New Brunswick, NJ: National Institute for Early Education Research.

Barnow, B.S., & Cain, G.G. (1977). A reanalysis of the effect of Head Start on cognitive development: Methodology and empirical findings. *The Journal of Human Resources, 12*(2), 177-197.

Burchinal, M.R., Lee, M.,& Ramey, C. (1989). Type of day-care and preschool intellectual development in disadvantaged children. *Child Development, 60*(1), 128-137.

Campbell, F.A., Pungello, E.P., Miller-Johnson, S., Burchinal, M., & Ramey, C.T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology, 37*(2), 231-242.

Currie J., & Thomas, D. (1995). Does Head Start make a difference? *The American Economic Review, 85*(3), 341-364.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics, 1*(3), 111-134.

Deutsch, M., Taleporos, E., & Victor, J. (1974). A brief synopsis of an initial enrichment program in early childhood. In S. Ryan (Ed*.), A report on longitudinal evaluations of preschool programs, Volume 1: Longitudinal evaluations (*pp. 49-60). Washington, DC: Office of Child Development, US Department of Health, Education, and Welfare.

Frede, E., Jung, K., Barnett, S.W., & Figueras, A. (2009). *The APPLES blossom: Abbott Preschool Program Longitudinal Effects Study (APPLES) preliminary results through 2nd grade*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.

Frede, E., Jung, K., Barnett, W.S., Lamy, C.E., & Figueras, A. (2007). *The Abbott Preschool Program longitudinal effects study (APPLES): Interim report.* New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.

Goodman, A., & Sianesi, B. (2005). Early education and children's outcomes: How long do the impacts last?. *Fiscal Studies, 26*(4), 513-548.

Studies Used in Test Score Fadeout Analysis (Continued):

Gormley W.T., Jr., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-k program. *The Journal of Human Resources, 40*(3), 533-558.

Gormley, W.T., Jr., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-k on cognitive development. *Developmental Psychology, 41*(6), 872-884.

Gormley, W.T., Jr., Phillips, D., & Gayer, T. (2008). Preschool programs can boost school readiness [Supplemental material]. *Science, 320*, 1723-1724. doi: 10.1126/science.1156019.

Heckman, J.J., Pinto, R., Shaikh, A.M., & Yavitz, A. (2011). *Inference with imperfect randomization: The case of the Perry Preschool program* (Working Paper No. 16935). Cambridge, MA: National Bureau of Economic Research.

Herzog, E., Newcomb, C.H., & Cisin, I.H. (1974). Double deprivation: The less they have, the less they learn. In S. Ryan (Ed.), *A report on longitudinal evaluations of preschool programs, Volume 1: Longitudinal evaluations* (pp. 69-94). Washington, DC: Office of Child Development, US Department of Health, Education, and Welfare.

Hustedt, J.T., Barnett, W.S., & Jung, K. (2008). *Longitudinal effects of the Arkansas Better Chance program: Findings from kindergarten and first grade.* New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.

Hustedt, J.T., Barnett, W.S., Jung, K., & Thomas, J. (2007). *The effects of the Arkansas Better Chance program on young children's school readiness.* New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.

Hustedt, J.T., Barnett, W.S., Jung, K., & Figueras-Daniel, A. (2009). *Continued impacts of New Mexico pre-k on children's readiness for kindergarten: Results from the third year of implementation*. New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.

Huston, A., Gupta, A., & Schexnayder, D. (2012). *Study of early education in Texas: The relationship of pre-K attendance to 3rd grade test results.* Austin, TX: University of Texas.

Jung, K., Barnett, W.S., Hustedt, J.T., & Francis, J. (2013). *Longitudinal effects of the Arkansas Better Chance program: Findings from first grade through fourth grade.* New Brunswick, NJ: Rutgers University, National Institute for Early Education Research.

Lee, V.E., Brooks-Gunn, J., & Schnur, E. (1988). Does Head Start work?: A 1-year follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Developmental Psychology, 24*(2), 210-222.

Lee, V.E., Brooks-Gunn, J., Schnur, E., & Liaw, F.R. (1990). Are Head Start effects sustained? A longitudinal follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Child Development, 61*(2), 495-507.

Lipsey, M.W., Hofer, K.G., Dong, N., Farran, D.C., & Bilbrey, C. (2013). *Evaluation of the Tennessee voluntary prekindergarten program: End of pre-K results from the randomized control trial.* Nashville, TN: Vanderbilt University, Peabody Research Institute.

Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R.W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review, 26*(1), 52-66.

Magnuson, K.A., Ruhm, C., & Waldfogel, J. (2007). The persistence of preschool effect: Do subsequent classroom experiences matter?. *Early Childhood Research Quarterly, 22*(1), 18-38.

Magnuson, K.A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review, 26*(1), 33-51.

Malofeeva, E., Daniel-Echols, M., & Xiang, Z. (2007). *Findings from the Michigan School Readiness Program 6 to 8 follow up study.* Ypsilanti, MI: High/Scope Educational Research Foundation.

Peisner-Feinberg, E.S., & Schaaf, J.M. (2010) *Long-term effects of the North Carolina More at Four pre-kindergarten program: Children's reading and math skills at third grade.* Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute.

Peisner-Feinberg, E.S., & Schaaf, J.M. (2011). *Evaluation of the North Carolina More at Four Pre-Kindergarten Program.* Chapel Hill, NC: University of North Carolina, FPG Child Development Institute.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., ... & Spier, E. (2010). *Head Start impact study: Final report.* Washington, DC: US Department of Health and Human Services.

Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., & Downer, J. (2012) *Third Grade Follow-Up to the Head Start Impact Study: Final Report* (OPRE Report 2012-45). Washington, DC: US Department of Health and Human Services.

Reynolds, A.J., & Temple, J.A. (1995). Quasi-experimental estimates of the effects of a preschool intervention. *Evaluation Review, 19*(4): 347-373.

Roy, A. (2003). *Evaluation of the Head Start Program: Additional evidence from the NLSCM79 data* (Doctoral dissertation, University at Albany, State University of New York).

Schweinhart, L.J., Barnes, H.V., & Weikart, D.P. (1993*). Significant benefits: The High/Scope Perry Preschool Study through age 27.* Ypsilanti, MI: High/Scope Press.

Sontag, M., Sella, A.P., & Thorndike, R.L. (1969). The effect of Head Start training on the cognitive growth of disadvantaged children. *The Journal of Educational Research, 62*(9), 387-389.

Studies Used in Test Score Fadeout Analysis (Continued):

Vance, B.J. (1967). The effect of preschool group experience on various language and social skills in disadvantaged children: Final Report. Stanford, CA: Stanford University.

Wasik, B.H., Ramey, C.T., Bryant, D.M., & Sparling, J.J. (1990) A longitudinal study of two early intervention strategies: Project CARE. *Child Development, 61*(6), 1682-1896.

Weiland, C., & Yoshikawa, H. (2013) Impacts of a prekindergarten program on children' mathematics, language, literacy, executive function, and emotional skills. *Child Development, 84*(6), 2112-2130.

Wong, V.C., Cook, T.D., Barnett, W.S., & Jung, K. (2008). An effectiveness-based evaluation of five state pre-kindergarten programs. *Journal of Policy Analysis and Management, 27*(1), 122-154.

Xiang, Z., & Schweinhart, L. J. (2002). *Effects five years later: The Michigan School Readiness Program evaluation through age 10.* Ypsilanti, MI: High/Scope Educational Research Foundation.

Zhai, F., Brooks-Gunn, J., & Waldfogel, J. (2011). Head start and urban children's school readiness: A birth cohort study in 18 cities. *Developmental Psychology, 47*(1), 134-152.

Zigler, E., Abelson, W.D., Trickett, P.K., & Seitz, V. (1982). Is an intervention program necessary in order to improve economically disadvantaged children's IQ scores? *Child Development, 53*(2), 340-348.

4.8 Valuation of Mental Health Outcomes

WSIPP'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, disruptive behavior, serious mental illness, and post-traumatic stress disorder. The category of disruptive behavior covers the DSM categories of oppositional defiant disorder and conduct disorder, while the category of serious mental illness includes DSM diagnoses such as major depression, schizophrenia, and bipolar disorder that result in serious functional impairment. Obviously, there are other recognized mental health disorders. It is anticipated that future development of WSIPP's model will include additional categories. This section of the Technical Documentation describes WSIPP's current procedures to estimate the monetary benefits of program-induced changes in these mental health conditions.

In general, WSIPP's mental health modeling follows the same analytic procedures described in Chapter 4.4 for alcohol, tobacco, and other drugs. Readers can refer to that section to find more detail.

WSIPP'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.[166] For example, Insel (2008) summarizes findings indicating the total cost of serious mental illness in the US in 2002 to be $317.6 billion in "economic" costs ($1,081 per capita) with 31.5% of this total due to health care expenditures, 60.8% due to loss in labor market earnings, and 7.7% percent due to disability payments.[167] 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 WSIPP'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 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 or 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.

---

[166] See, for example, Harwood, H., Ameen, A., Denmead, G., Englert, E., Fountain, D., & Livermore, G. (2000). 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. Heeringa, C., Lakoma, S., Petukhova, M.D., Rupp, M., Schoenbaum, A.E., . . . 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.

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

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 WSIPP'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. These costs include the costs of inpatient, outpatient, emergency, office-visit, and pharmacy services, excluding the costs of mental health treatment.
- 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 and serious mental illness.

4.8a Mental Health Parameters.

WSIPP'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. In addition, there are several other input parameters used in the mental health model that are general to WSIPP's overall benefit-cost model and these are discussed elsewhere in this chapter. In the following sections, the sources for the parameters and the computational routines are described.

4.8b Mental Health Epidemiological Parameters

WSIPP'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.

The methods used to compute the current prevalence of mental health conditions are the same as those used to compute the current prevalence of ATOD disorders; please see Chapter 4.4b for formulas and detailed descriptions.

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

Exhibit 4.69 displays the current parameters in WSIPP's model for the first three epidemiological factors, along with sources and notes. The death probability information is described in Section 4.8c in this Chapter and displayed in Exhibit 4.70.

Exhibit 4.69
Input Parameters for the Epidemiology of Mental Health Disorders

	DSM ADHD (a)	DSM depression (b)	DSM anxiety (c)	DSM serious mental illness (d)	DSM post-traumatic stress disorder (e)	Disruptive behavior (f)
Percent of population with lifetime DSM disorder[1]	8.1%	23.2%	31.5%	11.1%[4]	8.7%	9.0%
Age of onset						
Type of distribution[2]	Beta-	Beta-	Beta-general	Beta-	Beta-	Beta-
Parameter 1	17.362	1.1615	0.40667	1.046	1.248	1.8705
Parameter 2	41.582	2.1852	2.1615	2.153	2.669	1.2511
Parameter 3	3	9	5	10	3	3
Parameter 4	18	79	79	78	80	18
Persistence of DSM disorder, given onset						
Type of distribution[3]	Lognormal	Beta-	Beta-general	Beta-	Beta-	Lognormal
Parameter 1	3.2391	0.51946	0.82942	0.50164	0.61995	1.869
Parameter 2	1.5097	2.6936	2.0051	4.4312	1.7813	1.122
Parameter 3	n/a	0	0	0	0	n/a
Parameter 4	n/a	138.09	196.67	409.53	278.69	n/a

Notes and sources:

1) Except where noted: 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) *Ibid*. All age of onset distributions were fit with data reported in 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. For disruptive behavior, we combined the onset curves from oppositional defiant disorder and conduct disorder. For SMI, we combined the data for major depression and bipolar reported in Kessler et al. (2005) with an estimate for the onset of schizophrenia taken from Sham, P.C., MacLean, C.J., & Kendler, K.S. (1994). A typological model of schizophrenia based on age at onset, sex and familial morbidity. *Acta Psychiatrica Scandinavica,* 89(2), 135-41. We weighted the data by lifetime prevalence of each disorder in the population, then used @Risk software to estimate the best-fitting probability density distribution for SMI, as described above.
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 US 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. For SMI, we combined the survival curves from bipolar disorder and major depression with a hypothetical curve for schizophrenia (no remission). We weighted these curves by the prevalence of these disorders in the general population, and followed the procedure described above to estimate the persistence distribution for SMI.
4) Because the Kessler et al. (2005) paper does not provide lifetime prevalence estimates for serious mental illness, we computed the weighted ratio of estimated lifetime prevalence to 12-month prevalence of major depression and bipolar disorder from the NCS-R (Kessler et al., 2005). This ratio was then applied to the estimated 12-month prevalence of SMI from the National Survey on Drug Use and Health, 2012 (http://www.nimh.nih.gov/statistics/SMI_AASR.shtml) to approximate a lifetime prevalence of SMI.

4.8c Mental Health Attributable Deaths

WSIPP'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 a mental health disorder. The model inputs for these calculations are shown in Exhibit 4.70 below. For both of these disorders, we assume that a proportion of deaths by suicide are caused by a mental illness.

Exhibit 4.70

Mental Health Disorder-Annual Attributable Deaths by Age Group, 2006-2010

Age group	Years in age group	Number of suicides (all cases)	All deaths in state	State population in age group	% of suicides attributable to depression	% of suicides attributable to SMI
0-14	15	4	632	1,309,139	50%	25%
15-19	5	40	190	449,500	50%	25%
20-24	5	71	352	467,031	50%	25%
25-34	10	127	810	946,195	50%	25%
35-44	10	156	1,216	905,468	50%	25%
45-54	10	204	3,324	966,058	50%	25%
55-64	10	134	6,437	880,718	50%	25%
65-74	10	67	8,422	512,730	50%	25%
75-84	10	52	11,965	257,808	50%	25%
85-100	16	30	16,708	123,123	50%	25%

Depression. For suicides, the data source is the US Department of Health and Human Services, Centers for Disease Control (CDC). CDC estimates, for each state, the number of deaths attributable to suicide ("intentional self-harm"). The estimates from CDC are available online via a database called *WONDER*.[168] According to CDC:

> *The Underlying Cause of Death data available on WONDER are county-level national mortality and population data spanning the years 1999-2010. Data are based on death certificates for US residents. Each death certificate identifies a single underlying cause of death and demographic data.*

The CDC/ARDI estimates for Washington State are the average annual number of CDC/ARDI deaths, by age group shown in Exhibit 4.70, for the years 2006-10.

To compute depression-induced death rates for these age groups, we obtain 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 2006-10, the same years as the CDC death estimates. We assume that 50% of suicides are caused by depression.

Serious Mental Illness (SMI). As for depression, we use suicide death data from the CDC's *WONDER* database for the years 2006 to 2010. We assume that 25% of suicides are caused by serious mental illness (i.e., serious forms of depression, schizophrenia, or bipolar disorder). Although there is a large body of research about the increased mortality risk of people with serious mental illness,[169] many of the reported underlying causes of death are similar to the general population: lower socio-economic status SMI people with poor health and poor access to health care are more likely to die earlier than those who do not have those risk factors, similar to the general population.[170] Untangling the causal relationships between poor mental health and poor physical health may allow us to more accurately value mortality risk in seriously mentally ill populations in the future; at this stage, we only estimate the increased risk of suicide due to serious mental illness, not other causes of death.

---

[168] Centers for Disease Control and Prevention website: http://wonder.cdc.gov/

[169] Colton C.W., & Manderscheid, R.W., (2006). Congruencies in increased mortality rates, years of potential life lost, and causes of death among public mental health clients in eight states. *Prevention Chronic Disease, 3*(20), A42. Dembling, B.P., Chen, D.T., & Vachon, L. (1999). Life expectancy and causes of death in a population treated for serious mental illness. *Psychiatric Services (Washington, D.C.),* 50(8), 1036-42; Miller, B.J., Paschall, C.B. ., & Svendsen, D.P. (2006). Mortality and medical comorbidity among patients with serious mental illness. *Psychiatric Services (Washington, D.C.),* 57(10), 1482-7.; Parks, J., & National Association of State Mental Health Program Directors. (2006). *Morbidity and mortality in people with serious mental illness*. Alexandria, VA: National Association of State Mental Health Program Directors (NASMHPD) Medical Directors Council.

[170] Druss, B.G., Zhao, L., Von, E.S., Morrato, E.H., & Marcus, S.C. (2011). Understanding excess mortality in persons with mental illness: 17-year follow up of a nationally representative US survey. *Medical Care,* 49(6), 599-604.

For each type of mental illness, the death data are used to compute the probability of dying from the disorder in the general population, by age group, using the following equation:

$$(4.71) \quad MHD_a = \left((Suicides_a \times MHSuicidePct)/Pop_a\right)/Years_a$$

The probability of dying from a particular mental illness in each age group in the general population, MHD_a, is computed by multiplying the deaths due to suicide, $Suicide_a$, to the mental illness-specific proportion of suicides due to that disorder $MHSuicidePct$, 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. The value of death is monetized with the value of a statistical life described in Section 4.11d.

4.8d Linkages: Mental Health to Other Outcomes

WSIPP'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 WSIPP model are listed in the Appendix.

4.8e Human Capital Outcomes Affecting Labor Market Earnings via Mental Health Morbidity and Mortality

The WSIPP 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 Chapter 4.1, WSIPP's model uses national earnings data from the US Census Bureau's Current Population Survey (CPS). The CPS data used in this analysis represent average earnings of all people, both workers and non-workers at each age.

Exhibit 4.71

Labor Market Parameters for Mental Health Morbidity and Mortality

		Depression	Anxiety	Serious mental illness	PTSD
	Distribution Type	Gamma	Gamma	LogNormal	LogNormal
Gain in labor market earnings for never used vs. current disordered users, probability density distribution parameters	Alpha/Mean	48.2100	35.89100	-0.13176	-0.91337
	Beta/Std Dev.	0.00790	0.01331	0.16565	0.16825
	Shift	0.82202	0.77321	0.60002	0.77767
	Distribution Type	Gamma	Gamma	LogNormal	LogNormal
Gain in labor market earnings for former users vs current disordered users, probability density distribution parameters	Alpha/Mean	48.2100	35.89100	-0.13176	-0.91337
	Beta/Std Dev.	0.00790	0.01331	0.16565	0.16825
	Shift	0.82202	0.77321	0.60002	0.77767

Using the same methods as for ATOD, 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 but do not currently have a disorder. 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 the following equation:

$$(4.72) \quad EarnC_y = \frac{EarnAll_y \times (1 + EarnEscAll)^{y-tage} \times EarnBenAll \times (1 + EarnBenEscAll)^{y-tage} \times \left(IPD_{base}/IPD_{cps}\right)}{\left((1 + EarnGN) \times \left(1 - \left(CP_y + \left(\sum_{o=1}^{y}(O_o \times LTP) - CP_y\right)\right)\right) + (1 + EarnGF) \times \left(\sum_{o=1}^{y}(O_o \times LTP) - CP_y\right) + CP_y\right)}$$

The numerator in equation 4.72 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, *IPDcps*, with the year chosen for the overall analysis, *IPDbase*. These variables are described in Chapter 4.1.
The denominator in equation 4.72 uses the epidemiological variables described above: age of onset probabilities, *Oy,* lifetime prevalence rates, *LTP*, and current 12-month prevalence rates, *CPy*, 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 the Appendix.

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. The Appendix displays the results of our meta-analysis of these two bodies of research for DSM mental health disorders. Our meta-analytic procedures are described in Chapter 2.1.

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 2013 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 as shown in Chapter 4.1d. 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 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 a mental health 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 modeled. The distribution parameters are entered in the model, as shown in Exhibit 4.71. In the Monte Carlo analysis, we randomly draw probabilities as seeds for the modeled distribution. 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 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 the following equation:

$$(4.73)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta MH_y \times \left(1 - \sum_{o=1}^{y} O_o\right) \times EarnGN \times EarnC_y\right) + \left(\Delta MH_y \times \left(1 - \left(1 - \sum_{o=1}^{y} O_o\right)\right) \times EarnGF \times EarnC_y\right)}{(1 + dis)^{(y - tage + 1)}}$$

Where ΔMHy is the change in mental health disorder probability; *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 primarily people without a disorder, but may also serve some who have the 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 the following equation:

$$(4.74)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta MH_y \times EarnGF \times EarnC_y\right)}{(1 + dis)^{(y - tage + 1)}}$$

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.

We also model the change in expected labor market earnings due to mortality. The present value of future labor market earnings at each age is multiplied by the decrease in probability that a person dies as the result of the disorder given that they have the disorder at that particular age.

Valuing Employment for Individuals with Serious Mental Illness. For many intervention programs treating people with serious mental illness, the aim is to improve functioning of those individuals, not necessarily to relieve their mental illness itself. Therefore, we developed an alternative method of estimating labor market outcomes for populations with serious mental illness. For programs measuring employment in seriously mentally ill populations, we estimate the change in earnings caused by a program by multiplying the change in employment produced by the program by the expected earnings of a person with serious mental illness as shown in the following equation:

$$(4.75)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta Emp_y \times EarnSMI\right)}{(1+dis)^{(y-tage+1)}}$$

EarnSMI is estimated by multiplying the average earnings of workers for that age by a single factor, *PctSMIEarn*. This factor was estimated by comparing the average monthly earnings of Washington State Department of Social and Health clients with serious mental illness[171] with the average earnings of all workers from the CPS.

4.8f Medical Costs

WSIPP's model computes health care costs incurred (or avoided) with changes in the mental health conditions modeled. The inputs for these parameters are shown in Exhibit 4.72. They were computed from an analysis of data from the federal Medical Expenditure Panel Survey (MEPS).

Exhibit 4.72
Annual Expected Costs of Mental Health Conditions

		DSM ADHD	DSM depression	DSM anxiety	DSM serious mental illness	DSM post-traumatic stress disorder	Disruptive behavior
Child (age 1-17)	Annual $	599	537	537	0	537	599
	SE	270	290	290	0	290	270
	Year of $	2005	2005	2005	2005	2005	2005
Adult	Annual $	599	1,763	553	2,025	1,763	599
	SE	270	915	526	904	915	270
	Year of $	2005	2011	2011	2011	2011	2005

Estimates for mental health disorders. 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. The sample for MEPS includes approximately 15,000 individuals from the National Health Interview Survey. MEPS survey respondents in this subsample are followed over two years with five in-person interviews. In addition to documentation of medical encounters, the survey also provides information about demographics, family structure, comorbid conditions, insurance availability and other measures related to quality of life.

Indicators of mental health status in MEPS are only available for those individuals with a health care encounter. To estimate total health care related costs associated with a particular disorder, however, it is necessary to include individuals with the same condition who do not seek or receive treatment. The 2007 version of the NHIS was the most recent survey to ask adult respondents about the presence of mental health conditions. We identified adults with self-

[171] Data for fiscal year 2013 reported in the January 2014 Employment Monitoring Report for RSN Clients from DSHS Research and Data Analysis division, available here: http://www.dshs.wa.gov/rda/research/RSN/default.shtm

reported depression, anxiety, and serious mental illness (bipolar, psychosis, schizophrenia)[172] and linked these individuals to health care expenditure information from the 2008-2009 MEPS survey.[173] To estimate costs for adults with post-traumatic stress disorder, we used the results from depressed adults.

To assess mental health-related costs for children, we utilized data from the 2003 and 2004 version of the NHIS. These versions of the NHIS were the most recent year that included all 25 questions from the Strength and Difficulties Questionnaire (SDQ). The SDQ is a reliable and brief screening tool that rates the presence of four different psychological scales for children: emotional symptoms, conduct problems, hyperactivity/inattention, and peer relationship problems. The SDQ has been validated for children age four to 17. In each NHIS household, one sample adult and one sample child are randomly selected and additional questions are asked about this family member. The SDQ instrument is included in this "Sample Child Core" questionnaire. We used the "emotional symptoms" scale to estimate costs for "internalizing" problems: depression and anxiety in children, and the "hyperactivity/inattention" scale to estimate costs for "externalizing" problems: ADHD and disruptive behavior. Reponses for children in the sample child core questionnaire are linked to subsequent health care expenditures in the 2004-2005 MEPS survey.

There are two distinct challenges related to estimating the cost of health care attributable to a particular condition. The first challenge involves accounting for the likelihood that an individual will remain untreated (incur no costs). The second challenge stems from skewed data—a common occurrence in health care data when a small number of persons have excessive costs. To account for these issues, we developed two-part regression models following the methodology outlined in Glick, et. al.[174] The first part of the model predicts the (dichotomous) probability of incurring health care costs while the second part models the actual expenditure (conditional of receiving treatment). Our outcome variable of interest (expenditures) excluded treatment costs associated with mental illness (i.e. psychotherapy, antidepressants), but included other inpatient, outpatient, emergency room, office visit and pharmaceutical costs. Mental health related treatment costs were excluded since we were interested in potentially-avoidable health care costs that might be achieved with an effective intervention. Presumably, treatment related costs would persist following intervention as patients continued to manage their conditions. Regression models for each stage included the same set of covariates that might be expected to simultaneously correlate with mental illness and inflate total healthcare costs (e.g., age, presence of chronic illnesses, health insurance status, education, etc.).

The second part of this approach involved fitting the actual (untransformed) non-treatment expenditures using a generalized linear model (GLM). The two-part GLM allows for greater precision of estimated expenditures, compared to an ordinary least squares (OLS) regression with log transformed costs.[175] Different variance functions can be tested with a two-part GLM as well. To determine the best fitting functional family, we employed a modified Parks test,[176] which generally selected a Poisson distribution, reflecting the skewed nature of the data. Predicted expenditures are then obtained by multiplying the probability of having an expenditure (part one) by the estimated cost associated with the condition. Two expenditure estimates can be predicted from the model. First, we estimate the predicted expenditures for each person if we assumed the underlying disorder was present (and other characteristics remained constant). Then, using the same model, we estimate expenditures assuming the disorder was not present. Total expenditures attributable to the disorder equal the mean difference between these two estimates. All estimates were converted to 2012 dollars using Medical CPI.

[172] During the past 12 months, have you been frequently depressed?
During the past 12 months, have you been frequently anxious?
Have you ever been told by a doctor or other health professional that you had Bipolar Disorder? Mania or psychosis? Schizophrenia?

[173] http://www.cdc.gov/nchs/nhis/nhismep.htm

[174] Glick, H. (2007). *Economic evaluation in clinical trials*. Oxford: Oxford University Press

[175] Buntin, M.B., & Zaslavsky, A.M. (2004). Too much ado about two-part models and transformation?. *Journal of Health Economics,* 23(3), 525-542.

[176] Glick, Ibid.

Exhibit 4.73

Two-Part Model Assessing Non-Treatment Healthcare Costs of Adult Depression

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age (ref: 18-34)	35 to 44	-0.19		(-0.48 - 0.10)
	45 to 54	-0.35	**	(-0.66 - -0.05)
	55 to 64	0.14		(-0.32 - 0.60)
	65 to 74	0.61	*	(-0.03 - 1.26)
	75 and older	0.70	*	(0.00 - 1.41)
Female		0.96	***	(0.73 - 1.19)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.66	***	(-0.91 - -0.40)
	Black, non-Hispanic	-0.47	***	(-0.72 - -0.21)
	Asian, non-Hispanic	-0.61	***	(-1.01 - -0.22)
Marital status (ref: Married)	Widowed	-0.74	**	(-1.32 - -0.16)
	Divorced	-0.53	***	(-0.84 - -0.22)
	Never married	0.02		(-0.23 - 0.26)
Ever uninsured during year		-0.60	***	(-0.83 - -0.37)
Has usual source of medical care		1.19	***	(0.97 - 1.41)
Education (ref: Less than HS)	High school	0.41	***	(0.13 - 0.68)
	Some college or degree	0.49	***	(0.18 - 0.80)
Lives in metro area		-0.14		(-0.42 - 0.14)
Number of chronic conditions		0.82	***	(0.69 - 0.96)
Limitation in physical functioning		0.81	***	(0.31 - 1.32)
Self-reported depression (last year)		**0.12**		**(-0.30 - 0.54)**
Intercept		-0.03		(-0.48 - 0.43)
Part two: GLM, estimated costs				
Age (ref: 18-34)	35 to 44	-0.09		(-0.44 - 0.27)
	45 to 54	-0.05		(-0.42 - 0.32)
	55 to 64	0.45	*	(-0.04 - 0.94)
	65 to 74	0.40	*	(-0.05 - 0.86)
	75 and older	0.06		(-0.38 - 0.50)
Female		0.06		(-0.12 - 0.25)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.14		(-0.33 - 0.05)
	Black, non-Hispanic	0.02		(-0.16 - 0.21)
	Asian, non-Hispanic	-0.65	***	(-0.98 - -0.33)
Marital status (ref: Married)	Widowed	0.04		(-0.21 - 0.30)
	Divorced	-0.21		(-0.48 - 0.06)
	Never married	-0.20		(-0.43 - 0.04)
Ever uninsured during year		-0.54	***	(-0.82 - -0.26)
Has usual source of medical care		-0.06		(-0.36 - 0.25)
Education (ref: Less than HS)	High school	0.11		(-0.10 - 0.32)
	Some college or degree	0.21	*	(-0.02 - 0.45)
Lives in metro area		-0.13		(-0.49 - 0.23)
Number of chronic conditions		0.16	***	(0.07 - 0.25)
Limitation in physical functioning		0.75	***	(0.51 - 0.99)
Self-reported depression (last year)		**0.36**	**	(0.04 - 0.68)
Intercept		7.95	***	(7.34 - 8.55)

no. of obs = 5,522
weighted size = 229,038,154
Design df = 200
F(20, 181) = 31.78
Prob > F =0.0000

Exhibit 4.74

Two-Part Model Assessing Non-Treatment Healthcare Costs of Adult Serious Mental Illness (SMI)

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age (ref: 18-34)	35 to 44	-0.09		(-0.36 - 0.19)
	45 to 54	-0.21		(-0.53 - 0.11)
	55 to 64	0.24		(-0.20 - 0.68)
	65 to 74	0.77	**	(0.13 - 1.40)
	75 and older	0.82	**	(0.12 - 1.51)
Female		1.06	***	(0.85 - 1.27)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.69	***	(-0.96 - -0.43)
	Black, non-Hispanic	-0.49	***	(-0.73 - -0.25)
	Asian, non-Hispanic	-0.71	***	(-1.09 - -0.32)
Marital status (ref: Married)	Widowed	-0.80	***	(-1.37 - -0.22)
	Divorced	-0.59	***	(-0.90 - -0.28)
	Never married	-0.06		(-0.31 - 0.19)
Ever uninsured during year		-0.82	***	(-1.04 - -0.61)
Education (ref: less than HS)	High school	0.49	***	(0.22 - 0.76)
	Some college or degree	0.58	***	(0.28 - 0.88)
Lives in metro area		-0.01		(-0.27 - 0.26)
Census region (ref: West)	Midwest	0.05		(-0.29 - 0.39)
	Northeast	0.08		(-0.29 - 0.45)
	South	0.05		(-0.24 - 0.33)
Number of chronic conditions		0.91	***	(0.78 - 1.05)
Limitation in physical functioning		0.78	***	(0.28 - 1.28)
Told by doctor have serious mental health condition (lifetime)		0.66	***	(0.20 - 1.12)
Intercept		0.40		(-0.10 - 0.91)
Part two: GLM, estimated costs				
Age (ref: 18-34)	Age: 35 to 44	-0.01		(-0.26 - 0.23)
	Age: 45 to 54	0.01		(-0.28 - 0.31)
	Age: 55 to 64	0.46	***	(0.15 - 0.77)
	Age: 65 to 74	0.47	***	(0.16 - 0.79)
	Age: 75 and older	0.10		(-0.21 - 0.42)
Female		0.23	***	(0.07 - 0.39)
Race/ethnicity (ref: White, non-Hispanic)	Race/eth: Hispanic	-0.12		(-0.32 - 0.08)
	Race/eth: Black, non-Hispanic	0.08		(-0.12 - 0.28)
	Race/eth: Asian non-Hispanic	-0.72	***	(-1.01 - -0.43)
Marital status (ref: Married)	Marital: Widowed	0.03		(-0.21 - 0.27)
	Marital: Divorced	-0.24	**	(-0.43 - -0.04)
	Marital: Never married	-0.14		(-0.36 - 0.09)
Ever uninsured during year		-0.58	***	(-0.78 - -0.39)
Education (ref: Less than HS)	Education: High school	0.06		(-0.14 - 0.26)
	Education: Some college or degree	0.10		(-0.11 - 0.31)
Lives in metro area		-0.15		(-0.45 - 0.15)
Census region (ref :West)	Midwest	-0.15		(-0.37 - 0.07)
	Northeast	-0.23	**	(-0.44 - -0.02)
	South	-0.27	***	(-0.47 - -0.07)
Number of chronic conditions		0.18	***	(0.12 - 0.24)
Limitation in physical functioning		0.77	***	(0.59 - 0.95)
Told by doctor have serious mental health condition (lifetime)		0.37	**	(0.07 - 0.66)
Intercept		7.99	***	(7.51 - 8.47)

no. of obs = 5,522
weighted. size = 229,038,154
Design df = 200
F(22, 179) = 24.66
Prob > F = 0.0000

Exhibit 4.75

Two-Part Model Assessing Non-Treatment Healthcare Costs of Adult Anxiety Disorders

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age (ref: 18-34)	Age: 35 to 44	-0.22		(-0.92 - 0.47)
	Age: 45 to 54	-0.35		(-1.05 - 0.35)
	Age: 55 to 64	-0.56		(-1.26 - 0.15)
	Age: 65 to 74	-0.10		(-0.89 - 0.70)
	Age: 75 and older	0.19		(-0.67 - 1.06)
Female		1.04	***	(0.83 - 1.25)
Ever uninsured during year		-1.04	***	(-1.26 - -0.82)
Number of chronic conditions		0.85	***	(0.72 - 0.98)
Limitation in physical functioning		0.72	***	(0.24 - 1.19)
Census region (ref :West)	Midwest	0.24		(-0.07 - 0.54)
	Northeast	0.25		(-0.10 - 0.60)
	South	0.10		(-0.15 - 0.36)
Takes daily aspirin		0.84	***	(0.45 - 1.23)
Self-reported anxiety (last year)		**-0.02**		**(-0.39 - 0.35)**
Intercept		0.77	**	(0.10 - 1.43)
Part two: GLM, estimated costs				
Age (ref: 18-34)	Age: 35 to 44	-0.22		(-0.54 - 0.09)
	Age: 45 to 54	-0.18		(-0.47 - 0.11)
	Age: 55 to 64	-0.14		(-0.37 - 0.09)
	Age: 65 to 74	0.37	**	(0.03 - 0.71)
	Age: 75 and older	0.34	**	(0.08 - 0.59)
Female		0.24	***	(0.07 - 0.41)
Ever uninsured during year		-0.63	***	(-0.83 - -0.43)
Number of chronic conditions		0.15	***	(0.07 - 0.22)
Limitation in physical functioning		0.78	***	(0.60 - 0.97)
Census region (ref :West)	Midwest	-0.11		(-0.40 - 0.17)
	Northeast	-0.27	**	(-0.54 - 0.00)
	South	-0.24	*	(-0.51 - 0.02)
Takes daily aspirin		0.16	*	(-0.01 - 0.33)
Self-reported anxiety (last year)		**0.13**		**(-0.09 - 0.34)**
Intercept		8.02	***	(7.57 - 8.47)

no. of obs = 5,522
weighted. size = 229,038,154
Design df = 200
F(14, 187) = 35.27
Prob > F = 0.0000

Exhibit 4.76

Two-Part Model Assessing Non-Treatment Healthcare Costs of Child Emotional Conditions (Depression, Anxiety)

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age (ref:13 to 17)	Age: 4 to 12	0.32	***	(0.12 - 0.52)
Female		0.26	**	(0.06 - 0.45)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.80	***	(-1.04 - -0.57)
	Black, non-Hispanic	-0.83	***	(-1.13 - -0.52)
	Asian non-Hispanic	-0.72	***	(-1.17 - -0.26)
Number of chronic conditions		0.49	***	(0.19 - 0.78)
Poverty status (ref; high income)	Poor	-0.46	***	(-0.75 - -0.17)
	Low income	-0.51	***	(-0.80 - -0.21)
	Middle income	-0.35	**	(-0.63 - -0.06)
Ever uninsured during year		-0.76	***	(-1.02 - -0.51)
Emotional condition indicated (SDQ)		**0.17**		**(-0.24 - 0.59)**
Intercept		1.33	***	(1.08 - 1.58)
Part two: GLM, estimated costs				
Age (ref:13 to 17)	Age: 4 to 12	-0.37	**	(-0.66 - -0.08)
Female		0.40	**	(0.10 - 0.70)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.31	**	(-0.62 - -0.01)
	Black, non-Hispanic	-0.20		(-0.73 - 0.33)
	Asian non-Hispanic	-0.59	**	(-1.10 - -0.08)
Number of chronic conditions		0.87	***	(0.42 - 1.32)
Poverty status (ref; high income)	Poor	-0.14		(-0.48 - 0.21)
	Low income	-0.09		(-0.58 - 0.40)
	Middle income	0.28		(-0.07 - 0.64)
Ever uninsured during year		-0.53	***	(-0.85 - -0.22)
Emotional condition indicated (SDQ)		**0.49**	**	**(0.07 - 0.92)**
Intercept		6.76	***	(6.41 - 7.11)

no of obs = 3,390
weighted size = 111,376,243
Design df = 227
F(11, 217) = 13.83
Prob > F = 0.0000

Exhibit 4.77

Two-Part Model Assessing Non-Treatment Healthcare Costs of Child Hyperactivity-Related Conditions

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age (ref:13 to 17)	Age: 4 to 12	0.33	***	(0.13 - 0.53)
Female		0.26	**	(0.06 - 0.46)
Race/ethnicity (ref: White, non Hispanic)	Hispanic	-0.80	***	(-1.03 - -0.56)
	Black, non-Hispanic	-0.83	***	(-1.13 - -0.52)
	Asian non-Hispanic	-0.73	***	(-1.19 - -0.28)
Number of chronic conditions		0.51	***	(0.21 - 0.80)
Poverty status (ref; high income)	Poor	-0.43	***	(-0.72 - -0.14)
	Low income	-0.49	***	(-0.78 - -0.20)
	Middle income	-0.34	**	(-0.62 - -0.05)
Ever uninsured during year		-0.77	***	(-1.02 - -0.52)
Hyperactivity condition indicated (SDQ)		-0.24		(-0.58 - 0.10)
Intercept		1.34	***	(1.09 - 1.60)
Part two: GLM, estimated costs				
Age (ref:13 to 17)	Age: 4 to 12	-0.41	***	(-0.71 - -0.11)
Female		0.37	**	(0.07 - 0.68)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.32	**	(-0.63 - -0.02)
	Black, non-Hispanic	-0.17		(-0.71 - 0.36)
	Asian non-Hispanic	-0.60	**	(-1.11 - -0.09)
Number of chronic conditions		0.85	***	(0.40 - 1.30)
Poverty status (ref; high income)	Poor	-0.09		(-0.45 - 0.26)
	Low income	-0.06		(-0.54 - 0.42)
	Middle income	0.31	*	(-0.04 - 0.66)
Ever uninsured during year		-0.50	***	(-0.83 - -0.17)
Hyperactivity condition indicated (SDQ)		0.12		(-0.19 - 0.44)
Intercept		6.81	***	(6.45 - 7.18)

no of obs = 3,390
weighted size = 111,376,243
Design df = 227
F(11, 217) = 13.77
Prob > F = 0.0000

Valuing Specific Health Care Costs for Individuals with Serious Mental Illness. As described in the section on employment for seriously mentally ill individuals, intervention programs treating people with serious mental illness aim to improve functioning of those individuals, not necessarily to relieve their mental illness itself. Therefore, we developed an alternative method of estimating health care costs for populations with serious mental illness. For programs measuring the specific outcomes of psychiatric hospitalization, general hospitalization, or emergency department visits in seriously mentally ill populations, we estimate the change in health care costs caused by a program by multiplying the change in the specific outcome produced by the program by the expected cost of that outcome for a person with serious mental illness, as shown in the following equation:

$$(4.76) \quad PV\Delta HC = \sum_{y=tage}^{100} \frac{\left(\Delta HCOutcome_y \times HCCostSMI\right)}{(1+dis)^{(y-tage+1)}}$$

In formula 4.76, *HCCostSMI* is estimated from the sources listed in Exhibit 4.79. In addition, the expected change in outcome resulting from a program is based on an expected base rate of that outcome for a seriously mentally ill individual, based on the annual likelihood that a seriously mentally ill person will use that service. The cost and base rate inputs are displayed in Exhibit 4.78, the sources of these inputs are described in Exhibit 4.79.

Exhibit 4.78

Expected Costs of Health Care Resources Used by Individuals with Serious Mental Illness

	Emergency department	Hospital (general)	Hospital (psychiatric)
Annual $	1,448	15,082	21,356
SE	150	1,768	19,709
Year of $	2011	2011	2012
Annual percent of seriously mentally ill adults using resource	29%	20%	8%

Exhibit 4.79

Expected Annual Likelihood and Costs of Services for Individuals with Serious Mental Illness: Sources of Estimates

	Cost	Base rate
Emergency department visits	WSIPP analysis of 2007 MEPS data; sample-weighted average cost of ED visits for those classified as SMI, conditional on having at least one ED visit in the year.	WSIPP analysis of National Survey on Drug Use and Health, 2012: of those with past year SMI, proportion who were treated in the emergency room at least once in past year (NMERTMT2).
General hospitalization	WSIPP analysis of 2007 MEPS data; sample-weighted average cost of inpatient visits for those classified as SMI, conditional on having at least one inpatient visit in the year.	WSIPP analysis of National Survey on Drug Use and Health, 2012: of those with past year SMI, proportion who stayed in hospital overnight as inpatient in past year (INHOSPYR).
Psychiatric hospitalization	Weighted average of 2012 average cost of a psychiatric unit discharges from Washington State Comprehensive Hospital Abstract Reporting System (CHARS) system, and 2012 average cost of a client in the state mental hospitals, provided by DSHS Research and Data Analysis division.	Sum of 2012 psychiatric unit discharges from Washington State Comprehensive Hospital Abstract Reporting System (CHARS) system, and 2012 number of clients residing in the state mental hospitals, provided by DSHS Research and Data Analysis division, divided by estimated total population of seriously mentally ill individuals in Washington.

4.9. Valuation of Health Care Outcomes

WSIPP's benefit-cost model contains procedures to estimate the monetary value of 1) changes in health care service utilization, 2) changes in the incidence of certain health conditions, and 3) changes in total costs of care resulting from health care interventions The model estimates the value of programs designed to reduce hospitalizations, emergency room visits, cesarean sections, and total costs of care. The model also monetizes the outcomes of diabetes, obesity, and changes in weight or body mass index (BMI). Obviously, there are other health conditions and outcomes with economic consequences. It is anticipated that future development of WSIPP's model will include additional categories. This section of the Technical Documentation describes WSIPP's current procedures to estimate the monetary benefits of program-induced changes in these utilizations and decreases in incidence of health conditions.

The purpose of WSIPP'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 cost of care and health conditions. 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 health care resource utilization and in health conditions. If, for example, empirical evidence indicates that a primary care-based treatment program can reduce obesity, or reduce unnecessary visits to the emergency room, then what long-run benefits, if any, can be expected from these improved outcomes? 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.

Many programs and interventions measure utilization of specific health care resources. We have begun to cost out these resources for specific populations targeted by the interventions we have investigated so far (such as chronically ill individuals or new mothers receiving Medicaid). We have focused our efforts thus far on the outcomes that are most frequently reported in the research literature, and where we have been able to estimate costs. The utilization measures include hospitalization, hospital readmissions, emergency room visits, cesarean sections, and total costs of care.

WSIPP's modeling of health conditions follows the same general analytic procedures described in Chapter 4.8 for mental health disorders. Readers can refer to that section to find more detail. For example, to look at the long-term economic implications of the health conditions of diabetes and obesity, WSIPP's 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 health conditions.

We also provide economic estimates for other related outcomes when the health conditions of diabetes and obesity are not measured by outcome evaluations. For example, we examine the economic implications of weight loss through its causal link to diabetes (this relationship is discussed in Section 4.9f).

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

- Labor market earnings from health morbidity or mortality, to the degree there is evidence that current earnings are reduced because of health conditions (morbidity), or lifetime earnings are lost because of premature death (mortality) caused by health conditions.
- Health care costs for health morbidity, to the degree that these costs are caused by health conditions. These costs include the costs of inpatient, outpatient, emergency, office-visit, and pharmacy services.
- Value of a statistical life (VSL) estimates, net of labor market gains, applied to the change in mortality estimated to be caused by diabetes and obesity.
- Hospital admission, readmission, emergency department, and Caesarian section costs, to the degree that interventions (e.g., case management for frequent ED user, care transition programs) reduce utilization.
- Total costs of care, to the degree that interventions (e.g., patient-centered medical homes) reduce costs.

4.9a Health Care Parameters.

WSIPP's health care model is driven by three sets of parameters: 1) general parameters used throughout the benefit-cost model (shown in Exhibits 4.80 and 4.81); 2) parameters describing prevalence and cost of specific health care resources such as hospitalization and emergency department utilization (Exhibits 4.82-4.86); 3) inputs describing various aspects of each modeled health condition's epidemiology and linked relationships with other outcomes. Health conditions are addressed in Section 4.9b. First, we describe the general health care parameters.

Total Washington personal health care expenditures are collected from the Centers for Medicare & Medicaid Services, US Department of Health & Human Services. We use the ratio between "Drugs and other medical non-durables" and "Hospital care" to compute an added drug cost for every hospital visit we monetize. A hospital cost-to-charge ratio for Washington State is computed with 2011 data from the Healthcare Cost and Utilization Project (HCUP) of the US Department of Health & Human Services. Total annual emergency room visits in Washington for 2008 is computed from data compiled by the Washington State Hospital Association.

Exhibit 4.80

General Health Care Parameters

Parameter	Value
Total Washington personal health care expenditures[1]	$45,246,000,000
Hospital care	$16,074,000,000
Drugs and other medical non-durables	$5,386,000,000
Hospital cost-to-charge ratio[2]	0.3551
Emergency department cost-to-charge ratio[3]	1.0
Emergency department admissions, 2008[4]	1,754,047

[1] Centers for Medicare & Medicaid Services, *Health expenditures by state of residence, 1991-2009*. Retrieved April 7, 2014 from http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/res-tables.pdf.
[2] Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project: http://hcupnet.ahrq.gov/
[3] WSIPP assumption
[4] Washington State Hospital Association. (2010). *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.

The model uses Washington State values for the proportional sources of state, local, and federal funding for the different types of health care expenditures, described in Exhibit 4.81 below. We also compute an estimate of the long-run real escalation rate in per capita inflation-adjusted personal health care costs from the 2009-2019 forecast from Centers for Medicare & Medicaid Services, US Department of Health & Human Services.[177] The Washington State model currently uses the same inputs for all types of health care costs (low =0.005, modal =0.018, high =0.027), but the model allows separate estimates for each type of cost.

[177] 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

Exhibit 4.81

Proportion of Health Care Costs by Source

	Total cost by perspective			Taxpayer cost by payer		
	Participant	Taxpayer	Other	State	Local	Federal
General health care[1]	12.71%	48.32%	38.97%	9.74%	0.00%	90.26%
Emergency department[1]	8.07%	49.35%	42.57%	8.60%	0.00%	91.40%
Mental health costs[2]	1.10%	18.20%	80.80%	27.26%	0.00%	72.74%
ATOD treatment[3]	12.71%	48.32%	38.97%	45.79%	3.69%	50.51%
General hospital[2]	2.96%	44.95%	52.09%	10.06%	0.00%	89.94%
Drug/pharmacy[2]	21.80%	33.30%	44.90%	15.65%	0.00%	84.35%
Nursing home[4]	33.71%	3.91%	62.38%	41.42%	0.00%	58.58%
Obesity: Under age 65[5]	12.77%	58.98%	28.24%	28.20%	0.00%	71.80%
Obesity: Age 65 and over[5]	12.67%	17.31%	70.02%	2.49%	0.00%	97.51%
Diabetes: Under age 65[5]	11.53%	49.26%	39.21%	21.88%	0.00%	78.12%
Diabetes: Age 65 and over[5]	11.37%	15.61%	73.02%	3.53%	0.00%	96.47%

[1] WSIPP calculation from 2011 Medical Expenditure Panel Survey data.
[2] Cost by perspective retrieved from the Washington State Comprehensive Hospital Abstract Reporting System (CHARS) system, for 2012. Taxpayer costs by payer calculated from 2010 Medical Expenditure Panel Survey data, available at: http://meps.ahrq.gov/mepsweb/data_stats/quick_tables_results.jsp?component=1&subcomponent=0&year=-1&tableSeries=1&searchText=&searchMethod=1&Action=Search.
[3] Cost by perspective is the same as general health care above; taxpayer costs by payer calculated from Washington State Department of Social and Health Services report: "*Overview of Publicly Funded Services Substance Use Prevention, Treatment and Recovery,*" retrieved from: https://www.dshs.wa.gov/sites/default/files/BHSIA/dbh/documents/WASubstanceUseServicesOverview03-20-13.pdf
[4] Cost by perspective calculated from the National Nursing Home Survey 2004.
[5] WSIPP calculation from 2013 Medical Expenditure Panel Survey data.

Second, we describe the health care resource utilization measures (see Exhibits 4.82-4.86). As the WSIPP benefit-cost model continues to develop, we will expand our ability to model these outcomes (and others), in an increasing number of populations. To model the monetary benefits of reducing these utilization measures, we multiply the average cost of the measured resource for the specified population by the unit change produced from the program effect size and base rate for that population. For most resources, these are time-limited effects; e.g., reducing the likelihood of a hospitalization produces monetary benefits for a single year.

Exhibit 4.82

Hospitalization Parameters

	Children with asthma	Frequent emergency department users[1]
Average cost for a hospitalization[2]	$8,408	$18,563
Standard error on cost	$1,392	$4,761
Year of dollars	2012	2012
Annual likelihood of hospital admission[3]	3.5%	62.0%

[1] Frequent emergency department users are adults who visited the ED five times or more within a single year.
[2] WSIPP calculation from 2013 Medical Expenditure Panel Survey (MEPS) data.
[3] Of those in population, proportion who were admitted to hospital in a single year (MEPS).

Exhibit 4.83

Hospital Readmission Parameters

	Chronically ill adults[1]	General population	Women giving birth
Average cost for a readmission[2]	$19,642	$18,193	$6,600[3]
Standard error on cost	$3,453	$2,160	$2,422
Year of dollars	2012	2012	2009
Likelihood of readmission within 30 days after discharge[4]	24.8%	9.1%	2.0%[3]

[1] Chronically ill adults are those who are 1) at least 45 years old, and 2) have been diagnosed with one or more of the following conditions: coronary heart disease, angina, heart attack, other heart disease, diabetes, stroke, emphysema.
[2] WSIPP calculation from 2013 Medical Expenditure Panel Survey (MEPS) data.
[3] Weighted national estimates from a readmissions analysis file derived from the Healthcare Cost and Utilization Project (HCUP) State Inpatient Databases (SID), (2009), Agency for Healthcare Research and Quality (AHRQ). Accessed Dec 1, 2015 from: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb142.jsp.
[4] Of those in population and who had had at least one admission to the hospital, proportion who were re-admitted to hospital within 30 days of discharge (MEPS).

Exhibit 4.84

Emergency Department Parameters

	Children with asthma	Frequent emergency department users[1]	General population
Average cost for an ED visit[2]	$742	$4,641	$1,430
Standard error on cost	$65	$448	$35
Year of dollars	2012	2012	2012
Annual likelihood of ED visit[3]	18.8%	50.0%[4]	13.9%

[1] Frequent emergency department users are adults who visited the ED five times or more within a single year.
[2] WSIPP calculation from 2013 Medical Expenditure Panel Survey data (MEPS).
[3] Of those in population, proportion who were visited the emergency department in a single year (MEPS).
[4] Although this number is actually 100% (by definition), we use a 50% base rate for this population to maximize the unit change resulting from our effect size calculation.

Exhibit 4.85

Cesarean Section Parameters

	All mothers	Medicaid	Private-pay
Average cost for a cesarean section(compared to vaginal birth)[1]	$3,481	$3,021	$3,772
Standard error on cost	$121	$128	$178
Year of dollars	2014	2014	2014
Likelihood of cesarean section[2]	26.6%	24.0%	28.7%

[1] WSIPP analysis of pooled annual MEPS data from 2009-2013 period (five years). Expenditures have been converted to 2014 dollars using medical CPI.
[2] NTSV (primary) cesarean section rates in Washington State in 2008. From *Birth Statistics and Maternity Care Access*. (2010) Washington State Department of Social and Health Services--Planning, Performance, and Accountability Research and Data Analysis Division. Accessed Dec. 1, 2015 from: https://www.dshs.wa.gov/sites/default/files/SESA/rda/documents/research-9-98.pdf

Exhibit 4.86

Total health care cost parameters

	Chronically ill adults[1]	General population
Average annual cost for health care[2]	$7,235	$5,079
Standard deviation on cost	$16,717	$14,374
Year of dollars	2012	2012

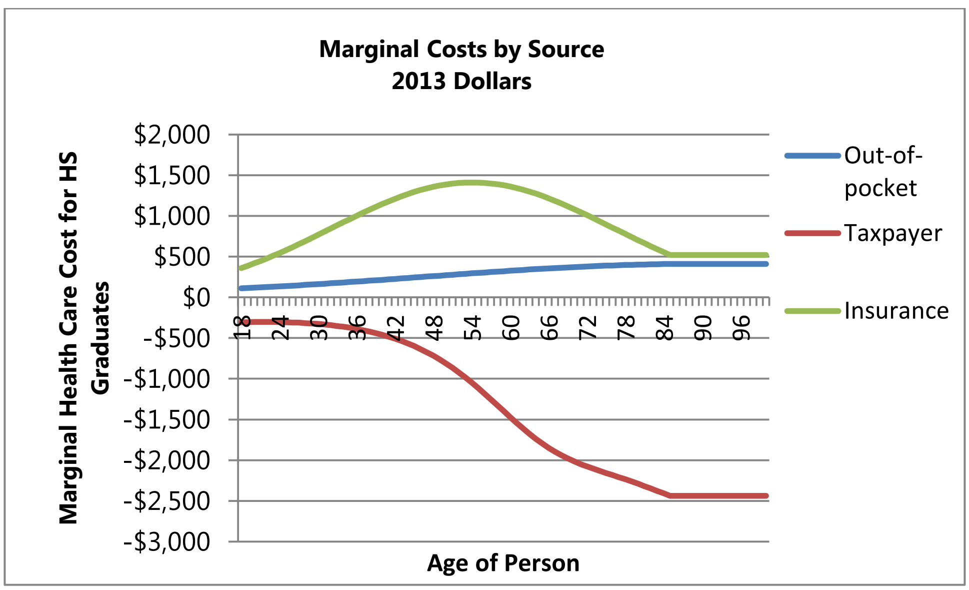
[1] Chronically ill adults are those who are 1) at least 45 years old, and 2) have been diagnosed with one or more of the following conditions: coronary heart disease, angina, heart attack, other heart disease, diabetes, stroke, emphysema.
[2] WSIPP calculation from 2013 Medical Expenditure Panel Survey data (MEPS).

Health Care Cost Estimates for High School Graduates Compared to those who do not Complete High School
An individual's level of education is related to health status and overall spending on healthcare over the course of a lifetime. Persons with higher levels of schooling are less likely to engage in risky health-related behaviors (smoking, excessive drinking) and more likely to exercise, obtain preventative health care (vaccines, mammograms, pap smears), and control chronic health conditions (diabetes, hypertension).[178] These differences, however, vary for individuals at different ages. To estimate health-care related costs and benefits over the lifespan, it is necessary to account for an individual's educational status.

For this analysis, we estimated differential health care costs by age for high school graduates and those without a high school diploma, after controlling for related factors. The analysis utilizes data from the 2013 Medical Expenditure Panel Survey (MEPS) which is based on a representative sample of the non-institutionalized population in the US. A two-part model—first examining the probability of having any health care expenditure, then estimating healthcare spending for those with costs—was developed for adult respondents in the MEPS data. The model controlled for demographic factors (age, sex, race/ethnicity), family characteristics (pregnancy, family size, marital status) and geographic region of residence (northeast, midwest, south, west). Healthcare expenditures included costs related to hospital inpatient care, hospital outpatient care, office-based medical provider services, emergency department services and prescriptions.

Since these cost differentials may differ by type of source of payment, we estimated two models for healthcare costs paid by private sources (individuals, health insurers) and public payers (Medicaid, Medicare, state programs). Based on these adjusted models, we calculated marginal estimates for the effect of high school graduation on healthcare costs at each single year of age from 18 to 85. Exhibit 4.87 shows the results of our analysis; overall, graduates tend to spend more out-of-pocket and through insurance, but less on publicly-funded health care resources.

Exhibit 4.87
Marginal Difference in Medical Costs by Education Level (High School Graduates vs. Non-Graduates)



4.9b Health Epidemiological Parameters

For the two health conditions currently modeled (obesity and diabetes), WSIPP's model begins by analyzing the epidemiology of each health condition 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.

[178] Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project: http://hcupnet.ahrq.gov/

The methods used to compute the current prevalence of health conditions are the same as those used to compute the current prevalence of ATOD disorders; please see Chapter 4.4b for formulas and detailed descriptions.

Four parameters enter the model to enable an estimate of the current prevalence of each health condition, from age one to age 100:

- Lifetime prevalence: the percentage of the population that has a specific health condition at some point during their lifetime;
- Age of onset: the age of onset of the specific health condition;
- Persistence: the persistence of the specific health condition, given onset; and
- Death (survival): the probability of death by age, after the age of treatment by a program.

Exhibit 4.88 displays the current parameters in WSIPP's model for the first three epidemiological factors, along with sources and notes. The death probability information is described in Section 4.9c in this Chapter.

In Exhibit 4.88, we provide parameter estimates for computing prevalence of diabetes and obesity for each age. Estimates for diabetes were derived from a variety of sources, described in the notes to Exhibit 4.88. Estimates for obesity were obtained using the NLSY—National Longitudinal Survey of Youth.[179] The NLSY included two cohorts of survey respondents. The 1979 cohort was made up of young women and men (ages 14-22) who were born between 1957 and 1964.[180] Individuals from this cohort were surveyed annually between 1979 and 1994 and on a biennial basis after 1994. At the latest interview (2012), survey respondents were over 50 years old. The 1997 cohort included respondents who were born between 1980 and 1984 and were ages 12-17 when first interviewed in 1997. The 1997 cohort has been surveyed annually in 15 rounds; the latest interviews took place in 2011-12, when respondents were approximately 32 years old.

In each NLSY interview, physical characteristics of the respondent were recorded, such as height and weight. We calculated a Body Mass Index (BMI) figure for each individual using the formula: weight (lb) / [height (in)]^2 x 703. To determine standardized BMI scores for children and adolescents age 20 or younger, we utilized 2000 Centers for Disease Control (CDC) growth charts (http://www.cdc.gov/growthcharts/cdc_charts.htm). Based on CDC classifications, youth with an age-adjusted BMI over the 85th percentile were considered overweight while those above the 95th percentile were classified as obese. For adults, a BMI above 25 was categorized as overweight and obese was defined as a BMI score above 30.

179 http://www.bls.gov/nls
180 https://www.nlsinfo.org/content/cohorts/NLSY79

Exhibit 4.88

Input Parameters for the Epidemiology of Health Conditions

	Type 2 diabetes	Obesity
Percent of population with condition at any point in lifetime	37%[1]	58.4%[2]
Percent of at-risk (pre-diabetic/overweight) population with condition at any point in lifetime	70%[3]	84.1%[4]
Age of onset		
Type of distribution	Beta-general[5]	Beta-general
Parameter 1	4.007	6.0533
Parameter 2	2.5662	1.7113
Parameter 3	17.953	-35.762
Parameter 4	83.205	57.202
Persistence of DSM disorder, given onset		
Type of distribution	Static[7]	Logarithmic[8]
Parameter 1	1.0	0.9834
Parameter 2	n/a	-0.215
Parameter 3	n/a	n/a
Parameter 4	n/a	n/a

Notes and sources:

1) Preston, S., Fishman, E., & Stokes, A. (2014). Lifetime probability of developing diabetes in the United States. University of Pennsylvania Population Studies Center, PSC Working Paper Series, WPS 14-4. The estimate for the lifetime probability of developing diabetes is for the 1940-49 birth cohort taken from Table 1.
2) Among the 1979 NLSY cohort, 17.8% had become obese at some point prior to age 32, and 39.0% reached obesity prior to age 54. The incidence of obesity increased considerably among the more recent 1997 NLSY cohort. By age 32, 37.2% of this cohort had become obese at some point in their lifetime. We conservatively estimated that an additional 21.2% (*39.0% - 17.8% = 21.2%*) of the 1997 cohort would become obese by age 54 to derive our lifetime prevalence of 58.4%.
3) Recent studies suggest that 70% of individuals with prediabetes eventually develop the disease. See: Tabak A., Herder C., Rathmann W., Brunner, E., & Kivimaki, M. (2012). Prediabetes: a high-risk state for diabetes development. *The Lancet, 379,* 2279-2290; Perreault, L., Pan, Q., Mather, K., Waston, K., Hamman, R., & Kahn, S., (2012). Effect of regression from prediabetes to normal glucose regulation on long-term reduction in diabetes risk: results from the Diabetes Prevention Program Outcomes Study. *The Lancet, 379,* 2243-2251; and Gillett, M., Royle, M., Snaith, A., Scotland, G., Poobalan, A., Imamura, M., Black, C., Boroujerdi, M., Jick, S., Wyness, L., McNamee, P., Brennan, A., & Waugh, N. (2012). Non-pharmacological interventions to reduce the risk of diabetes in people with impaired glucose regulation: a systematic review and economic evaluation. *Health Technology Assessment, 16*(33), ISSN 1366-5278.
4) For youth who began the survey overweight in the 1997 NLSY, 84.1% became obese at some point prior to age 32. By comparison, only 60.4% of overweight individuals in the 1979 NLSY cohort became obese by age 32 and 82% of overweight individuals were obese by age 54. We retained our original estimate (84.1%) because we were not able to evaluate the obesity trajectory for overweight individuals in the 1997 cohort using historical trends.
5) Using @Risk software, we fit a probability density function to the estimates of annual diabetes incidence by age group (with no differential mortality), presented in Appendix 5 of Fishman, E.I., Stokes, A., & Preston, S.H. (2014). The dynamics of diabetes among birth cohorts in the US *Diabetes Care,* 37(4), 1052-1059.
6) We combined data from two sources: Cunningham, S.A., Venkat, N.K.M., & Kramer, M.R. (2014). Incidence of childhood obesity in the United States. *New England Journal of Medicine,* 370(5), 403-411, and the National Longitudinal Survey of Youth. We recorded annual hazard rates of becoming obese for those who were normal weight at baseline, then created a cumulative distribution and normalized that distribution to 1. We then used @Risk software to fit a probability density function to the cumulative distribution.
7) We assume no remission from diabetes; this assumption is supported by Karter, A.J., Nundy, S., Parker, M.M., Moffet, H.H., & Huang, E.S. (2014). Incidence of remission in adults with type 2 diabetes: The diabetes & aging study *Diabetes Care, 37*(12), 3188-3195. The authors analyzed longitudinal data from over 120,000 Type-2 diabetic members of a health care system and found that only six maintained remission from diabetes for a period of five years or more, indicating essentially zero recovery from diabetes.
8) Persistence estimates for obesity are generated from cox proportional hazards models that predict obesity duration at given age ranges. Our final models examine obesity over a nearly thirty year period starting at age 20. The cohort that entered the National Longitudinal Survey of Youth (NLSY) in 1979 provides the most complete history for obesity patterns and forms the starting point for the analysis. In recent years, however, rates of obesity have increased substantially among younger adults. To account for the prevalence of obesity in more recent cohorts, we plotted known persistence curves for the youth entering the NLSY in 1997. Then, we generated predicted obesity duration estimates assuming this cohort followed a similar trajectory as the older (1979) cohort in later years. Estimated persistence probabilities are calculated at each year of age using the 'baseline' option in the proportional hazards regression (PHREG) procedure available in SAS 9.4.

4.9c Deaths Attributable to Health Conditions

WSIPP'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 a health disorder.

Diabetes. To estimate the proportion of deaths caused by diabetes, we relied on the work of Saydah et al.[181] The authors used data from the Second National Health and Nutrition Examination Survey (NHANES II) including its mortality component. The authors estimated a population attributable risk of death (for participants with diagnosed and undiagnosed diabetes, aged 30 to 74 at baseline) of 5.1%. We apply this diabetes-attributable death probability to all deaths in Washington.

Obesity. We used two rigorous studies to estimate the relative risk of death in obese individuals compared to those of normal weight.[182] Both studies controlled for smoking, a potential confounder, and underlying disease, a potential source of reverse causation. Calle et al., 1999 analyzed the mortality rates in a prospective cohort of 457,785 men and 588,369 women over 45 years old who were followed for 14 years. Using data from the NHANES, Calle et al., 2005 analyzed data the mortality rate in 317,875 men and women over 20 years. We computed a weighted average of the results from these studies and found a relative risk of death 1.5 times higher in individuals with a BMI over 30 kg/m^2 compared to individuals with a BMI of 23.5-24.9 kg/m^2.

For each type of health condition, the death data are used to compute the probability of dying from the disorder in the general population. We divide by the number of years in each age group to compute the annual probability of dying from the health condition among the general population. The value of the death is monetized with the value of a statistical life described in
Section 4.11d.

$$(4.77)\quad DiabDR_y = \frac{\frac{DiabD_a}{Pop_a}}{Years_a}$$

4.9d Linkages: Health Conditions to Other Outcomes

WSIPP's benefit-cost model monetizes improvements in health outcomes, in part, with linkages between each 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 diabetes and entering a nursing home 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 parameters are entered into the benefit-cost model and used when performing a Monte Carlo simulation. The linkages in the current WSIPP model are listed in the Appendix.

4.9e Human Capital Outcomes Affecting Labor Market Earnings via Health Condition Morbidity and Mortality

The WSIPP model computes lost labor market earnings as a result of 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 health condition (like diabetes or obesity). As described in Chapter 4.1, WSIPP's model uses national earnings data from the US Census Bureau's Current Population Survey (CPS). The CPS data used in this analysis represent average earnings of all people, both workers and non-workers at each age.

[181] Table 3, Saydah, S.H., Eberhardt, M.S., Loria, C.M., & Brancati, F.L. (2002). Age and the burden of death attributable to diabetes in the United States. *American Journal of Epidemiology, 156*(8), 714-719.

[182] Calle, E.E., Thun, M.J., Petrelli, J.M., Rodriguez, C., & Heath Jr, C.W. (1999). Body-mass index and mortality in a prospective cohort of US adults. *New England Journal of Medicine, 341*(15), 1097-1105.; Calle, E.E., Teras, L.R., & Thun, M.J. (2005). Obesity and mortality. *New England Journal of Medicine, 353*(20), 2197-2199.

Exhibit 4.89
Labor Market Parameters for Health Morbidity and Mortality

		Type 2 diabetes	Obesity
	Distribution type	Gamma	Gamma
Gain in labor market earnings for never used vs. current disordered users, probability density distribution parameters	Alpha	59.0410	60.5310
	Beta	0.00713	0.00554
	Shift	0.75942	0.73437
	Distribution type	Gamma	Gamma
Gain in labor market earnings for former users vs current disordered users, probability density distribution parameters	Alpha	59.0410	60.5310
	Beta	0.00713	0.00554
	Shift	0.75942	0.73437

Using the same methods as for mental health, for each person at each age, total CPS earnings can be viewed as a weighted sum of people who have never had a specific health condition, plus those that are currently in the condition, plus those that were formerly, but not currently in the condition (recovered). From the CPS data on total earnings for all people, the earnings of individuals with a current health condition, at each age, y, is computed with this equation:

$$(4.78)\quad EarnC_y = \frac{EarnAll_y \times (1+EarnEscAll)^{y-tage} \times EarnBenAll \times (1+EarnBenEscAll)^{y-tage} \times \left(IPD_{base}/IPD_{cps}\right)}{\left((1+EarnGN) \times \left(1-\left(CP_y + \left(\sum_{o=1}^{y}(O_o \times LTP) - CP_y\right)\right)\right) + (1+EarnGF) \times \left(\sum_{o=1}^{y}(O_o \times LTP) - CP_y\right) + CP_y\right)}$$

The numerator in equation 4.78 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, $IPDcps$, with the year chosen for the overall analysis, $IPDbase$. These variables are described in Chapter 4.1.

The denominator in equation 4.78 uses the epidemiological variables described above: age of onset probabilities, Oy, lifetime prevalence rates, LTP, and current 12-month prevalence rates, CPy, at each age.

The denominator also includes two variables on the earnings gain of people who have never had the health condition compared to those who currently have that condition, $EarnGN$, and the earnings gain of people who have recovered from the condition compared to those who currently have that condition, $EarnGF$. These two central relationships measure the effect of a 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 the Appendix.

For health conditions, just as for mental health disorders and ATOD, we meta-analyzed two sets of research studies. One set examines the relationship between health conditions and employment rates, and the second examines the relationship between health conditions and earnings, conditional on being employed. The Appendix displays the results of our meta-analysis of these two bodies of research for health conditions. Our meta-analytic procedures are described in Chapter 2.1.

For a health condition, from these two findings—the effect of a condition on employment, and the effect of a condition on the earnings of those employed—we then combine the results to estimate the relationship between a health condition 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 2011 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 individuals without the health condition to those with the condition 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 a health condition 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 chosen. The distribution parameters are entered in the model as shown in Exhibit 4.89. In the Monte Carlo analysis, we randomly draw probabilities as seeds for the modeled distribution. Since the body of evidence we reviewed in the meta-analysis did not allow separation of the effects into 1) people who never had the condition vs. those who currently have the condition, and 2) people who have recovered from the condition vs. those who currently have the condition, we enter the same normal 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 health condition is given by:

$$(4.79)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta H_y \times \left(1 - \sum_{o=1}^{y} O_o\right) \times EarnGN \times EarnC_y\right) + \left(\Delta H_y \times \left(1 - \left(1 - \sum_{o=1}^{y} O_o\right)\right) \times EarnGF \times EarnC_y\right)}{(1 + dis)^{(y - tage + 1)}}$$

Where ΔHy is the change in health condition probability; O are the annual onset probabilities; *EarnGN* is the earnings gain of people who never had the condition compared to people currently in the condition; *EarnGF* is the earnings gain of people who used to have the condition compared to those who currently have the condition; *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 condition and with a condition, 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 health condition is given by:

$$(4.80)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta H_y \times EarnGF \times EarnC_y\right)}{(1 + dis)^{(y - tage + 1)}}$$

This model for a treatment program is simpler than that for a prevention program because, by definition, a treatment program only attempts to turn people with a current condition into people who have recovered from that condition.

We also model the change in expected labor market earnings due to mortality. The present value of future labor market earnings at each age is multiplied by the decrease in probability that a person dies as the result of the disorder given that they have the disorder at that particular age.

4.9f Medical Costs for Specific Health Conditions

Exhibit 4.90 displays WSIPP's estimates for the total annual medical costs of diabetes and obesity, above and beyond what is observed in the general population of non-diabetic and non-obese individuals. Sources and methods for these estimates are described below.

Exhibit 4.90
Input Parameters for the Incremental Medical Costs of Health Conditions

	Type 2 diabetes	Obesity
Annual incremental cost of disorder	$2,418	$290
Standard error on annual cost	$344.85	$26.13
Year of dollars	2012	2014
Age at which cost was measured	47	18
Additional cost per year of life beyond measurement age	$29.47	$51.55
Standard error on additional cost	$6.27	$4.64

For health conditions like diabetes and obesity, WSIPP's approach to benefit-cost analysis models the incremental costs incurred (or avoided) with the inception (or reduction) of particular health care conditions. The cost of illness includes those expenditures directly associated with a condition as well as indirect costs that may be attributed to the presence of an underlying disease or disorder. Patients with certain health conditions (such as arthritis or bronchitis), for example, may experience chronic pain. However, expenses associated pain treatment may be related to multiple underlying conditions.

To estimate the total healthcare costs related to a condition, we follow the approach of Glick, et al., (2007) and estimate a two part model. The details of this approach are presented Section 4.8f. In short, the first part of the model accounts for the probability of having any health care expenditure among those diagnosed with a particular condition. The second stage models actual health care costs for those reporting expenditures. The adjusted estimates provide a realistic indication of the costs of a given condition, after accounting for utilization and other relevant factors.

Unless otherwise stated, cost of illness models are based on public data available in the federal Medical Expenditure Panel Survey (MEPS). Additional information about MEPS is provided in Section 4.8f. MEPS data are widely used in estimating health care costs, since this survey provides a comprehensive record of patient health encounters and accurate accounting of the payments associated with each visit or billed expense. While MEPS has notable strengths for this type of analysis, there are some limitations to this data. Generally, those limitations are related to information collected for a particular condition or population. The sections below discuss the condition-specific models and note any differences in our approach for each analysis.

Estimates for Diabetes. Diabetes represents one of the fastest growing health conditions in the US. In 2012, over 22.3 million Americans were diagnosed with diabetes (7% of the US population) compared with 17.5 million reported diabetics in 2007. According to the American Diabetes Association, total economic costs associated with diabetes exceeded $245 billion in 2012 and age-adjusted health care costs for diabetics were 2.3 times higher than costs for non-diabetics.[183] We utilized the 2012 MEPS household survey to identify individuals with a diabetes diagnosis and determine diabetes-related expenditures. Diabetes is listed as one of the "priority conditions" in the MEPS questionnaire—each person (age 18 or older) is asked if they were ever told by a doctor or health professional that they have diabetes.

Adults that self-reported a diagnosis of diabetes were provided a supplementary questionnaire called the Diabetes Care Survey (DCS). The DCS asked a series of questions about the respondent's diabetes, including age of onset, related symptoms (i.e. vision problems), use of insulin, and other diabetes management strategies.[184] In a small number of cases, the initial self-reported diabetes diagnosis is ruled out. Based on information provided in the 2012 DCS, we determined that 8.2% of all adults had a diagnosis of diabetes. Exhibit 4.91 shows the results for our two-part model of health care

[183] American Diabetes Association. (2013). Economic costs of diabetes in the US in 2012. *Diabetes Care,* 36(4), 1033-46.
[184] See http://meps.ahrq.gov/survey_comp/hc_survey/paper_quest/2012/2012_DCS_ENG.pdf

expenses related to diabetes. After accounting for the effects of gender, age, race/ethnicity and presence of other chronic health conditions, we estimate the annual health care expenses associated a 47 year old with diabetes was $2,418 (95% C.I. $1,741-$3,184). Using these model results, we applied an age-based escalator which adjusted this base cost by $29 for each year of age to account for differences in health care costs among younger/older diabetics.

Exhibit 4.91

Two-Part Model Assessing Healthcare Costs of Adult Diabetes

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Female		0.95	***	(0.87-1.03)
Age		0.01	***	(0.01-0.01)
Race/ethnicity (ref: Caucasian, non-Hispanic)	Hispanic	-0.68	***	(-0.81- -0.56)
	African American, non-Hispanic	-0.70	***	(-0.84- -0.56)
	Other race, non-Hispanic	-0.67	***	(-0.83- -0.50)
Insurance (ref: Uninsured)	Private	1.45	***	(1.33-1.5)
	Public	1.29	***	(1.16-1.42)
Chronic condition—arthritis		0.92	***	(0.74-1.10)
Chronic condition—asthma		0.85	***	(0.66-1.04)
Chronic condition—high blood pressure		0.90	***	(0.76-1.04)
Chronic condition—coronary heart disease		0.56	**	(0.08-1.04)
Chronic condition—cholesterol		0.87	***	(0.71-1.03)
Chronic condition—cancer		0.93	***	(0.56-1.30)
Chronic condition—emphysema		1.04	**	(0.15-1.93)
Chronic condition—diabetes		**1.39**	*******	**(1.07-1.72)**
Intercept		-0.61	***	(-0.77--0.45)
Part two: GLM, estimated costs				
Female		0.04		(-0.07 - 0.15)
Age		0.01	***	(0.00 - 0.01)
Race/ethnicity (ref: caucasian, non-Hispanic)	Hispanic	-0.07		(-0.24 - 0.09)
	African American, non-Hispanic	-0.12	**	(-0.23 - -0.02)
	Other race, non-Hispanic	-0.02		(-0.18 - 0.14)
Insurance (ref: Uninsured)	Private	0.76	***	(0.61 - 0.91)
	Public	0.80	***	(0.64 - 0.95)
Chronic condition—arthritis		0.57	***	(0.45 - 0.69)
Chronic condition—asthma		0.14	**	(0.02 - 0.25)
Chronic condition—high blood pressure		0.17	***	(0.05 - 0.28)
Chronic condition—coronary heart disease		0.38	***	(0.26 - 0.50)
Chronic condition—cholesterol		-0.10	*	(-0.20 - 0.01)
Chronic condition—cancer		0.48	***	(0.31 - 0.65)
Chronic condition—emphysema		0.30	***	(0.08 - 0.52)
Chronic condition—diabetes		**0.36**	*******	**(0.26 - 0.46)**
Intercept		7.21	***	(7.01 - 7.41)

no. of obs = 38,974
weighted size = 313,489,853
Design df = 203
F(15,189) = 170.30
Prob > F = 0.0000

Estimates for Obesity. We were unable to estimate the incremental annual health care costs for obese versus non-obese adults from the MEPS dataset. Instead, we computed a weighted average of annual cost estimates from seven high quality studies.[185] Average annual medical costs are estimated to be $290 (in 2014 dollars) higher for obese adults at age 18, compared to non-obese adults. These studies estimate the relationship between body mass index (BMI) and medical costs, controlling for gender, race, education, age, census region, household income, smoking status, and insurance status. More recent studies use instrumental variable estimation to account for the potential endogeneity of BMI. The effect of obesity on medical costs increases with age. The model allows for this by using the age profile of obesity-related costs estimated by An, (2015). Using data from An, we estimated that after age 18, average annual costs of obesity increased by an additional $52 per year of age. We also derived a coefficient of variation from An's findings and applied that to both the baseline annual cost at age 18 and the incremental cost by year of age to model the error in these estimates.

The Medical Expenditure Panel Survey[186] was used to estimate total health utilization and costs for populations of interest. Healthcare related expenses in MEPS include direct payments from individuals and payments from both public and private insurance. Expenditure estimates are obtained in a two-step process. First, survey respondents report healthcare utilization and expenditures for the household. Second, third-party payment information is obtained through a follow-back survey of health care providers to review related claims (called the Medical Provider Component (MPC)). Expenditure information includes both doctor and facility costs and is included in the MEPS Household Component (HC) file. The expenditure categories include emergency department, inpatient, and total health expenditures. Inpatient costs encompass all expenses for direct hospital care (room & board, diagnostic and laboratory work, x-rays and physician services). The total cost of health care includes expenses for medical provider (office); hospital care (outpatient, emergency department, and inpatient); prescribed medicine; home health; dental; and other medical expenses such as medical equipment and supplies, orthopedics, eye care, ambulance. Weighted estimates are based on expenditures reported in the 2012 MEPS. In some cases, multi-year estimates (2011-12) were utilized to account for small sample sizes in the population of interest.

Estimates for Diabetes Costs for Nursing Home Residents. Unfortunately, MEPS survey respondents do not include adults living in institutional facilities, such as nursing homes. According to the US Department of Health and Human Services (DHHS), 5.4% of the population age 75 or older lived in a nursing home in 2013. Given that the prevalence of diagnosed diabetes among this age group (75+) was approximately 23%, it is important to capture health related costs for those living in skilled nursing facilities as well. [187]

Exhibit 4.92 displays the assumptions and estimated annual costs we use when computing nursing home costs.

---

[185] An, R. (2015). Health care expenses in relation to obesity and smoking among US adults by gender, race/ethnicity, and age group: 1998-2011. *Public Health*, *129*, 29-36; Arterburn, D., Maciejewski, M., & Tsevat, J. (2005). Impact of morbid obesity on medical expenditures in adults. *International Journal of Obesity*, *29*, 334-339; Cawley, J., Meyerhoefer, C., Biener, A., Hammer, M., & Wintfeld, N. (2014). Savings in medical expenditures associated with reductions in body mass index among US adults with obesity, by diabetes status. *PharmacoEconomics,* [Epub ahead of print]; Wang, G., Zheng, Z., Heath, G., Macera, Cl, Pratt, M., & Buchner, D. (2002). Economic burden of cardiovascular disease associated with excess body weight in US adults. *American Journal of Preventive Medicine*, *23*, 1-6; Finkelstein, E., Fiebelkorn, I., & Wang, G. (2003). National medical spending attributable to overweight and obesity: How much, and who's paying? *Health Affairs*, W3:219-226; Finkelstein, E., Trogdon, J., Cohen, J., & Dietz, W. (2009). Annual medical spending attributable to obesity: Payer- and service-specific estimates. *Health Affairs, 28*(5), w822-w831; Baker, C., & Bradley, R. (2013). The simultaneous effects of obesity, insurance choice, and medical visit choice on healthcare costs. A chapter in *Measuring and Modeling Health Care Costs.* NBER [accessed at http://www.nber.org/chapters/c13118.pdf].

[186] http://meps.ahrq.gov

[187] Preston, S.H., Fishman, E., & Stokes, A., (2014). *Lifetime probability of developing diabetes in the United States.* PSC Working Paper Series, WPS 14-4.

Exhibit 4.92

Input Parameters for the Incremental Medical Costs of Health Conditions

	For nursing home residents
Annual cost of nursing home	$92,345
High annual cost	$132,053
Low annual cost	$36,938
Year of dollars	2014
Base rate of general population (age 75+) living in nursing home	5.4%
Age to begin costs	75

We obtained annual per-resident nursing home expenditures using the 2014 Genworth Cost of Care Survey for Washington State.[188] According to this survey, the median intermediate cost for a semi-private room was $253 dollars per day, or $92,345 per year (range $36,500-$132,300). Of course, the costs associated with diabetes represent only part of the total care costs in these facilities. We examined available research to determine the extent to which a diabetes diagnosis was related to nursing home admission. (See Appendix Exhibits 9.0 to 9.2 for a summary of the link between diabetes and nursing home utilization later in life.) The model attributes a portion of nursing home admission costs to diabetes incidence.

The estimates of health care expenditures obtained using MEPS data are apportioned according to primary payer. That is, costs are allocated to those borne by individuals, public payers (federal and state government) and private insurers. Since nursing home expenditures were not available in MEPS, we examined payments using the National Nursing Home Survey (NNHS).[189] The NNHS is a nationally-representative survey of 13,507 residents in 1,174 facilities that was last conducted in 2004. This step was important because one-third (33.7%) of nursing home costs are paid by individuals, compared to 11% for individuals living in the community. State-related Medicaid payments are also proportionally higher for nursing home residents compared to community-dwelling seniors (25.8% vs. 2.6%).[190]

4.10 Valuation of Labor Market Outcomes

This section describes WSIPP's benefit-cost modeling of labor market outcomes that are measured directly in program evaluations, not estimated via educational attainment, health condition, mental health disorder, or substance use disorder. Evaluations of programs such as workforce training strategies often measure the percent change in earnings for participants as a result of their participation in the program. Sometimes evaluations also measure changes in employment rates.

Earnings. The benefit-cost model directly monetizes changes to labor market earnings. The percent increase in earnings as a result of participation in a program is multiplied by the projected earnings for the specified population in each year (see Section 4.1d for a description of these populations). After adjusting for loss of earnings due to death in the participating population, the percent change is applied to earnings estimated using the parameters for the specific populations in equation 4.3.

Employment. Some programs do not measure changes in earnings directly. In such situations, we monetize the employment rate instead, which requires an extra step and assumption. We estimate the change in earnings caused by a program by multiplying the change in employment produced by the program by the expected earnings of a person as shown in the following equation:

[188] Genworth Financial., & National Eldercare Referral Systems. (CareScout). (2014). Genworth Financial 2014 cost of care survey: Home care providers, adult day health care facilities, assisted living facilities and nursing homes. Richmond, Va.: Genworth Financial.
[189] ICPSR04651-v1. Hyattsville, MD: US Dept. of Health and Human Services, National Center for Health Statistics [producer], 2004. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-03-23. http://doi.org/10.3886/ICPSR04651.v1
[190] See http://www.cdc.gov/nchs/data/nnhsd/2004NNHS_Use_of_Payment_Source_Data.pdf

$$(4.81)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{(\Delta Emp_y \times PopEarn)}{(1+dis)^{(y-tage+1)}}$$

PopEarn is estimated by dividing the expected earnings of the population analyzed by the percent of the population that is employed. Because of this extra step required in monetizing employment, we prefer the direct measure of labor market earnings, and use that where available.

4.11 Other Parameters

In addition to the parameters discussed in the previous sections of this Chapter, the model uses a number of additional user-supplied inputs to compute benefits and costs. These are discussed in this section.

4.11a 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, WSIPP converts dollars to the year specified by the user (currently 2015). 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 4.11f).

4.11b 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 4.82 indicates that the net present value of a program, evaluated at the at the age of a person for whom an investment is made, *NPVage*, is the discounted sum of benefits at each year, *By*, minus program costs at each year, *Cy*, discounted with a discount rate, *Dis*.

$$(4.82)\quad NPV_{age} = \sum_{y=age}^{N} \frac{B_y - C_y}{(1+Dis)^y}$$

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 low, modal, and high discount rates defining the triangle. Exhibit 4.93 shows the three discount rates are entered. WSIPP uses a low real discount rate of 2%, a modal rate of 3.5%, and a high rate of 5%. These input choices reflect the recommended rates in Moore et al., (2004).[191] Similarly, the Congressional Budget Office has used a 3% real discount rate in its analyses of Social Security.[192] 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% rate to summarize their main benefit-cost results.[193]

Exhibit 4.93
Discount Rates Used in Benefit-Cost Model

Range	Discount rate
Low value	0.020
Modal value	0.035
High value	0.050

[191] 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.
[192] Congressional Budget Office. (2012). *The 2012 Long-Term Projections for Social Security: Additional Information*. Washington, DC: Author. Retrieved August 8, 2013 from http://www.cbo.gov/sites/default/files/cbofiles/attachments/43648-SocialSecurity.pdf
[193] Heckman et al., (2010).

4.11c 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 4.94 displays a table with these inputs. For Washington State, we enter the current distribution of the state population by single year of age total state population from the Washington State Office of Financial Management (OFM), the official forecasting agency for the state. In addition, 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 produced by the US Department of Health and Human Services Centers for Disease Control and Prevention.[194] Since OFM does not break out population by year of age after the age of 85, WSIPP applies the CDC death expectancy rate to the to the previous year's population to estimate the population for those ages.

[194] Arias, E. (2015). *United States life tables, 2011* (National Vital Statistics Reports vol. 64, no. 11). Washington, DC: United States Department of Health and Human Services, National Vital Statistics System, Table 1.

Exhibit 4.94

Age	State population by single year of age for 2015	Life table, US (CDC, 2011)		Value of a statistical public cost year (2011 dollars)	
		Of 100,000, number surviving to age	Years of expected life remaining	Average medical payments	Average social security payments
1	87,950	100,000	78.7	1,605	0
2	86,882	99,394	78.2	604	0
3	88,210	99,353	77.2	389	0
4	88,870	99,327	76.2	294	0
5	88,419	99,306	75.3	311	0
6	89,598	99,289	74.3	339	0
7	90,489	99,274	73.3	463	0
8	92,220	99,261	72.3	784	0
9	91,801	99,249	71.3	550	0
10	89,081	99,239	70.3	383	0
11	89,037	99,230	69.3	406	0
12	88,160	99,221	68.3	319	0
13	86,975	99,212	67.3	391	0
14	86,967	99,201	66.3	461	0
15	89,333	99,184	65.3	1,370	0
16	89,944	99,159	64.4	625	0
17	89,444	99,127	63.4	554	0
18	89,365	99,087	62.4	331	0
19	89,416	99,039	61.4	390	0
20	90,105	98,982	60.5	318	0
21	92,030	98,917	59.5	458	0
22	94,387	98,844	58.6	445	0
23	97,756	98,762	57.6	342	0
24	98,647	98,675	56.7	299	0
25	98,314	98,584	55.7	418	0
26	98,358	98,493	54.8	3,468	0
27	96,054	98,400	53.8	331	0
28	94,774	98,308	52.9	566	0
29	94,750	98,214	51.9	492	0
30	97,788	98,117	51.0	1,311	0
31	99,164	98,017	50.0	839	0
32	97,802	97,914	49.1	545	0
33	99,908	97,807	48.1	716	0
34	100,011	97,696	47.2	361	0
35	99,027	97,582	46.2	489	0
36	99,819	97,465	45.3	751	0
37	92,137	97,343	44.3	422	0
38	91,670	97,215	43.4	367	0
39	89,753	97,080	42.5	835	0
40	88,069	96,937	41.5	538	0
41	89,997	96,784	40.6	739	0
42	85,213	96,621	39.6	460	0
43	86,998	96,446	38.7	627	0
44	91,480	96,256	37.8	822	0
45	98,533	96,047	36.9	992	0
46	99,635	95,816	36.0	1,025	0
47	93,149	95,561	35.1	573	0
48	90,035	95,283	34.2	982	0
49	88,223	94,979	33.3	970	0
50	89,297	94,645	32.4	955	0
51	95,627	94,281	31.5	1,390	0
52	96,970	93,884	30.6	710	0
53	98,936	93,456	29.8	2,447	0
54	98,220	92,995	28.9	1,443	0
55	100,045	92,502	28.0	943	0
56	100,448	91,975	27.2	1,885	0
57	97,066	91,411	26.3	1,769	0

Exhibit 4.94 (Continued)

Age	State population by single year of age for 2015	Life table, US (CDC, 2011)		Value of a statistical public cost year (2011 dollars)	
		Of 100,000, number surviving to age	Years of expected life remaining	Average medical payments	Average social security payments
58	98,355	90,808	25.6	1,618	0
59	96,495	90,164	24.7	1,362	0
60	95,265	89,476	23.9	1,197	0
61	95,006	88,746	23.1	2,451	0
62	91,132	87,971	22.3	1,435	0
63	89,712	87,150	21.5	2,157	0
64	84,380	86,280	20.7	2,220	0
65	81,045	85,355	20.0	1,631	0
66	79,990	84,368	19.2	3,620	14,657
67	76,755	83,306	18.4	4,591	14,657
68	77,763	82,160	17.7	6,610	14,657
69	75,423	80,923	16.9	8,436	14,657
70	56,511	79,597	16.2	5,237	14,657
71	56,887	78,184	15.5	7,558	14,657
72	55,251	76,675	14.8	7,373	14,657
73	52,994	75,062	14.1	6,497	14,657
74	45,302	73,335	13.4	5,595	14,657
75	40,865	71,489	12.7	5,344	14,657
76	37,912	69,513	12.1	8,244	14,657
77	35,320	67,405	11.5	4,887	14,657
78	33,383	65,160	10.8	7,295	14,657
79	30,285	62,766	10.2	9,974	14,657
80	28,225	60,212	9.6	10,055	14,657
81	26,688	57,493	9.1	4,708	14,657
82	23,872	54,614	8.5	7,864	14,657
83	22,750	51,593	8.0	8,063	14,657
84	21,461	48,434	7.5	5,465	14,657
85	20,711	45,145	7.0	7,920	14,657
86	18,981	41,733	6.5	8,047	14,657
87	17,237	38,200	6.1	8,047	14,657
88	15,654	34,593	5.7	8,047	14,657
89	14,073	30,955	5.3	8,047	14,657
90	12,513	27,338	4.9	8,047	14,657
91	10,992	23,798	4.6	8,047	14,657
92	9,529	20,389	4.2	8,047	14,657
93	8,141	17,169	3.9	8,047	14,657
94	6,845	14,187	3.7	8,047	14,657
95	5,656	11,486	3.4	8,047	14,657
96	4,587	9,097	3.2	8,047	14,657
97	3,643	7,036	3.0	8,047	14,657
98	2,830	5,307	2.8	8,047	14,657
99	2,146	3,897	2.6	8,047	14,657
100	1,584	2,782	2.4	8,047	14,657

4.11d Valuation of Reductions in Mortality Risk: Value of a Statistical Life

Several of the outcomes analyzed in WSIPP's benefit-cost model affect the risk of mortality. For example, as described in Chapter 4.4, 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.[195]

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 Chapter, WSIPP'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).[196] 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.[197] 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.[198]

WSIPP'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 avoided costs of health care[199] and Social Security[200] if someone dies. We also subtract the "human capital"

[195] For a general review of the analytical methods economists and others have used to assess the valuation of mortality risk, see W.I. Viscusi. (2008). *How to value a life* (Vanderbilt Law and Economics Research Paper No. 08-16), Nashville, TN: Vanderbilt University, Department of Economics.
[196] 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.
[197] 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.
[198] 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.
[199] To estimate health care costs by age for the average person in the population, we used data from the 2007 Medical Expenditure Panel Survey (MEPS), 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, by age of person receiving those services, was computed for 2007. These figures are adjusted for inflation and escalation in healthcare costs over time.
[200] We use an average Social Security payment of $14,657 in 2015 dollars, from age 66 on (Monthly Statistical Snapshot, April 2015, Social Security Administration). We escalate these dollars in future years using a 1.061% real growth rate, derived from "Annual Scheduled Benefit Amounts for Retired Workers With Various Pre-Retirement Earnings Patterns Based on Intermediate Assumptions," Social Security 2015 Trustees Report, https://www.ssa.gov/oact/TR/2015/V_C_prog.html#1078526.

derived benefits of changes to lifetime earnings (LTE) and household production (HP), described elsewhere in this document. Thus, the general approach is given in the following equation:

$$(4.83) \quad VSLAdj = VSL - HC - SS - LTE - HP$$

WSIPP's VSL model is driven with the parameters shown in Exhibit 4.95, along with life table and public cost year information displayed in Exhibit 4.94. The model includes 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).[201] 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.

The value of a statistical life year, $VSLY$, is then computed for the range of years considered in the Kniesner study (ages 18 to 62) with equation 4.84, 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 4.94.

$$(4.84) \; VSLY = \frac{disrate \times VSL}{1 - (1 + disrate)^{-L}}$$

Exhibit 4.95

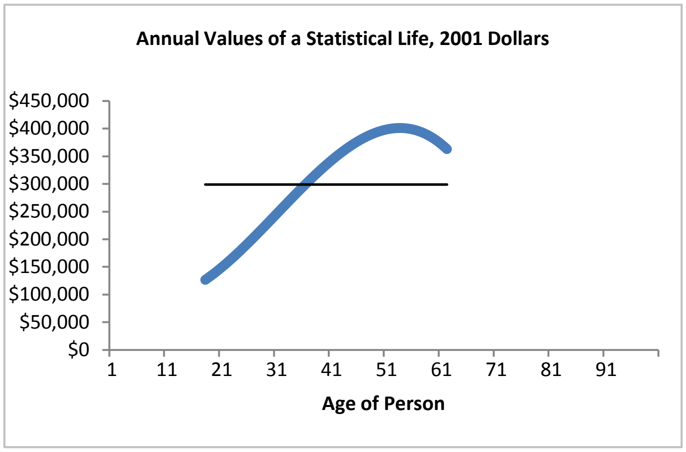
Value of a Statistical Life Parameters

Parameter	Value
Modal value of statistical life, millions	7.0
High value of statistical life, millions	10.0
Low value of statistical life, millions	4.0
Year of dollars	2001
Regression Parameter: Intercept	132.23
Regression Parameter: Age	-9.63
$Age^{\wedge 2}$	0.65
$Age^{\wedge 3}$	-0.007
Post-age 62 exponential change rate	-0.010
Pre-age 18 multiplier	1.70

For example, with a $7 million VSL (in 2001 dollars), a 3% discount rate, and 41 years of remaining life, the $VSLY$ is $299,000 on average over the ages of 18 to 62. The next set of parameters in Exhibit 4.94 are used to distribute this average $VSLY$ 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 above). The Aldy and Viscusi analysis, using revealed preference data from labor market wages, estimates the annual $VSLY$ for ages 18 to 62. Thus, by applying the third order polynomial to the base value ($299,000) the following distributed estimates of $VSLY$ are obtained for ages 18 to 62.

[201] 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.

Exhibit 4.96



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 Hersch[202] (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 4.95 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:

$$(4.85)\quad VSLY_y = VSLY_{62} \times (1 + esc)^{(y-62+1)}$$

For ages less that 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 compute 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 one to 18, $VSLY$ is computed with the Hammitt and Haninger ratio ($HHratio$):

$$(4.86)\quad VSLY_y = VSLY \times HHratio$$

4.11e 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.[203] 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 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 "Indirect 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%, 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.

$$(4.87)\quad DWL_{age} = \sum_{y=age}^{N} \frac{(B_y - C_y) \times DWL\%}{(1 + Dis)^y}$$

WSIPP uses a low real deadweight cost value of 0%, a modal rate of 50%, and a high rate of 100%. 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

[202] Viscusi, W.K., & Hersch, J. (2008). The mortality cost to smokers. *Journal of Health Economics, 27*(4), 943-958.

[203] 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.

program.[204] Also following Heckman et al. (2010), we do not apply any deadweight cost calculations to estimated taxes obtained from earnings related outcomes.[205]

4.11f 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 4.97 displays the price indices used by the model. The general inflation index used by WSIPP is the Implicit Price Deflator for Personal Consumption Expenditures produced by the Bureau of Economic Analysis. Since health care costs are central in WSIPP's benefit-cost model, and since health care prices have followed different paths than general prices, we also include a medical cost index. We use the BEA Implicit Price Deflator for Personal Consumption Expenditures for Health Services.

[204] Heckman et al., (2010).

[205] Ibid, see section J of the Heckman Appendix.

Exhibit 4.97

State Population and Inflation Indices by Year

Year	State population	Implicit price deflator[1]	IPD for health services[2]
1971	3,436,299	0.233	10.953
1972	3,430,299	0.241	11.418
1973	3,444,299	0.254	11.957
1974	3,508,700	0.280	13.093
1975	3,567,901	0.303	14.565
1976	3,634,904	0.320	15.969
1977	3,715,400	0.341	17.300
1978	3,836,199	0.365	18.760
1979	3,979,199	0.397	20.589
1980	4,132,156	0.440	23.082
1981	4,229,278	0.478	25.901
1982	4,276,549	0.505	28.816
1983	4,307,247	0.526	31.452
1984	4,354,067	0.546	33.806
1985	4,415,785	0.566	35.856
1986	4,462,212	0.578	37.988
1987	4,527,098	0.596	40.432
1988	4,616,886	0.620	43.584
1989	4,728,077	0.646	47.416
1990	4,866,692	0.674	51.462
1991	5,021,339	0.696	55.439
1992	5,141,178	0.715	59.336
1993	5,265,691	0.733	62.507
1994	5,364,342	0.748	65.047
1995	5,470,108	0.764	67.420
1996	5,567,764	0.780	69.007
1997	5,663,763	0.793	70.389
1998	5,750,030	0.799	71.716
1999	5,830,833	0.811	73.290
2000	5,894,143	0.831	75.354
2001	5,970,330	0.847	77.882
2002	6,059,316	0.859	79.883
2003	6,126,885	0.876	82.918
2004	6,208,515	0.897	85.984
2005	6,298,816	0.923	88.718
2006	6,420,258	0.947	91.410
2007	6,525,086	0.971	94.790
2008	6,608,245	1.001	97.352
2009	6,672,159	1.000	100.000
2010	6,724,540	1.017	102.517
2011	6,767,900	1.041	104.400
2012	6,817,770	1.061	106.305
2013	6,882,400	1.076	107.847
2014	6,968,170	1.091	109.063
2015	7061410	1.094	109.832

[1] Implicit Price Deflator-Personal Consumption Expenditures from Bureau of Economic Analysis, National income and Product Account Tables. Table 1.1.9 Implicit Price deflators for Gross Domestic Product, Line 2. Accessed from: http://www.bea.gov/iTable/iTable.cfm?reqid=9&step=3&isuri=1&903=13#reqid=9&step=3&isuri=1&904=1971&903=13&906=a&905=1000&910=x&911=0, April 8, 2016.

[2] Implicit Price Deflator-Personal Consumption Expenditures for Health Services. Bureau of Economic Analysis, National income and Product Account Tables. Table 2.3.4 Price Indexes for Personal Consumption Expenditures by Major Type of Product, Line 16. Accessed from: http://www.bea.gov/iTable/iTable.cfm?ReqID=9#reqid=9&step=3&isuri=1&904=1971&903=64&906=a&905=2016&910=x&911=0, April 8, 2016

4.11g Tax Rates

The benefit-cost model uses average tax rates for several calculations. We used the aggregate total from the Tax Foundation from 2015 to represent a combination of all kinds of taxes paid (income, sales, property, and other), as a percentage of income.[206] This value and the breakdown are displayed in Exhibit 4.98.

Exhibit 4.98
Tax Rates

	Percent of total, by source		
Total tax rate	**Federal**	**State**	**Local**
0.3123	0.6492	0.2044	0.1465

In addition, we allow the user to input the ultimate sources of the tax rate, i.e., what proportion of taxes paid go to state, local, and federal sources. We follow the procedures of the Tax Policy Center to break down Government receipts and expenditures as reported in the Bureau of Economic Analysis National Income and Product Accounts Tables for those parameters.[207]

4.11h Capital Costs

A few routines in the model use capital financing costs. The real cost of capital of 0.05 was obtained from discussions with fiscal staff of the Washington State Legislature.

4.12 Value of an Outcome

The WSIPP benefit-cost model is used to evaluate the incremental effects of programs and policies. For example, if an education policy increased the chance of high school graduation from 70% to 75%, the model monetizes the gains from that improvement. The model can also be used to estimate the "full effect" of an outcome. For example, we can compare the monetary value of someone who graduated from high school to someone who does not. We call these larger effects a "value of an outcome" calculation.

The value of an outcome calculations are useful in that they allow us to compare out estimates to those made by other researchers. There are bodies of research, for example, on the lifetime value of high school graduation, the lifetime cost of child abuse and neglect, the lifetime costs of diabetes, and so on. By comparing our results to those of other researchers, we can determine the degree to which our model aligns with the best studies that have focused on a given topic.

Exhibit 4.99 provides a brief overview of some comparisons made between WSIPP benefit-cost values and those of other researchers, for several of the outcomes evaluated in the WSSIP model.

To try to make our computations comparable to others', we adjust a few parameters in our model to match those used by another researcher. We adjust the year of the dollars to match that used by the other researcher and the discount rate to match that used by the researcher. Additionally, some researchers only consider a subset of the ways we monetize outcomes in the WSIPP model, and sometimes the other researchers include more ways to monetize outcomes that we do.

---

[206] We looked at data from two separate sources: 1) Pomerleau, Kyle. (2015). *Tax Freedom Day® 2015*. Washington, DC: Tax Foundation. Retrieved May 18, 2016 from: http://taxfoundation.org/sites/taxfoundation.org/files/docs/TaxFoundation_TFD_Report.pdf. 2) Citizens for Tax Justice (2015,). Who pays taxes in America in 2015? Washington, DC: author. Retrieved August 9, 2013 from: http://www.ctj.org/pdf/taxday2015.pdf. The first source gave a federal estimate of a total effective tax rate of 31.2%, while the second source gave an estimate of 30.0%. Because these numbers were so similar, we used the Tax Foundation number of 31.2%.

[207] To breakdown total government receipts between federal, state, and local sources, we used the methods from the Tax Policy Center (a collaboration between the Urban Institute and the Brookings Institution). Retrieved May 18, 2016 from: http://www.taxpolicycenter.org/briefing-book/background/numbers/revenue-breakdown.cfm. Bureau of Economic Analysis tables can be found at. U.S. Bureau of Economic Analysis, "Section 3 – Government Current Receipts and Expenditures," http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=1&isuri=1&903=1 (accessed April 5, 2016).

Exhibit 4.99

WSIPP's Benefit-Cost Values Compared to other Researchers

Outcome	Comparison study			WSIPP benefit-cost model		Common year of dollars and discount rate
	Comparison study	Key result	Notes on study	WSIPP result	Comparison note	
Cigarette Smoking, Total Lifetime Costs	Frank A. Sloan, Jan Ostermann, Gabriel Picone, Christopher Conover and Donald H. Taylor, Jr. The Price of Smoking. MIT Press: 2004.	$170,789	From author's table 11.4. The analysis is estimated for a 24 year old smoker. Their number with equal males and females is $162,975.	$115,724	We adjusted WSIPP input parameters (for year in which dollars are denominated, the discount rate, and the age of the person) to match Sloan's. The comparison with Sloan may be apples to oranges because we currently model persistence of the 24 year old, and it is not clear that he does this (except for death).	Year 2000 dollars, 3% discount rate
Cigarette Smoking, Annual Health Care Cost	An, R. (January 01, 2015). Health care expenses in relation to obesity and smoking among US adults by gender, race/ethnicity, and age group: 1998-2011. Public Health, 129, 1, 29-36.	$1,046	MEPS and NHIS for smokers and non-smokers 18 and older.	723.29 (358.91 base + incremental by year cost of 7.84 per year)	We also use a MEPS and NHIS based national number. We classify our programs into preventing smoking or stopping smoking. This comparison is to the number for treatment.	Year 2011 dollars
High School Graduation, Labor Market Earnings	Rouse, Cecilia Elena. Consequences for the Labor Market Chapter in *The Price We Pay. 2007*. Editors Belfield, Clive R., Levin, Henry M.	$190,230 if just HS $386,392 if continue on to more education at rate of hs grads	Ages 20-67. 2004 dollars. Uses Cross-Sectional Differences in CPS. GEDs treated as HS graduates, excludes prison pop, military.	$278,898 from LME w/externality and cost of Higher Ed $329,687 w/ externality and no cost of Higher Ed	We adjusted WSIPP input parameters (for year in which dollars are denominated, the discount rate, and began program and effect at age 18,). GEDs and late graduations are not treated as graduates. Author does not use a causality factor and we use one from Heckman by education. We do use labor market gains and costs from continuing on to further education.	Year 2004 dollars, 3.5% discount rate, 0% productivity/earnings/benefits growth

Exhibit 4.99 (Continued)

WSIPP's Benefit-Cost Values Compared to other Researchers

Outcome	Comparison study			WSIPP benefit-cost model		Common year of dollars and discount rate
	Comparison study	Key result	Notes on study	WSIPP result	Comparison note	
High School Graduation, Total Social Value	Belfield, Hollands, Levin. Providing Comprehensive Education Opportunity to Low-Income Students: What are the Social and Economic Returns	$415,700 in Labor Market Earnings, $542,261 overall	Ages 18-64. NY-based projection. The Belfield estimate does not include the gateway effect. It only compares HSGrad to HSDropouts.	$273,989 Labor Market Earnings (including externality), $280,122 Overall	Assuming that students become HSGrad but do not continue on to further education.	Year 2011 dollars, 1% productivity/earnings growth, 0% benefit growth, 3.5% discount rate
Child Abuse and Neglect	Fang, X., Brown, D. S., Florence, C. S., & Mercy, J. A. (February 01, 2012). Economic Burden of Child Maltreatment in the US and Implications for Prevention. Child Abuse & Neglect, 36, 2.)	$210,012	For a 6 year old.	$199,684	We use a lower labor market escalation rate than Fang	Year 2010 dollars, 3% discount rate
Obesity, Total Lifetime Costs	Kasman, M., Hammond, R., Werman, A., Mack-Crane, A., & McKinnon, R. (2015) An In-Depth Look at the Lifetime Economic Cost of Obesity [PowerPoint slides]. Retrieved from http://www.brookings.edu/~/media/events/2015/05/12-economic-costs-of-obesity/0512-obesity-presentation-v6-rm.pdf	$92,235	Focusing on ages 25-85.	$99,381	Started at age 25 and extended through modeled life.	Year 2013 dollars, 3% discount rate
Diabetes, Lifetime Health Care Cost	Zhuo, X., Zhang, P., Barker, L., Albright, A., Thompson, T. J., & Gregg, E. (January 01, 2014). The lifetime cost of diabetes and its implications for diabetes prevention. Diabetes Care, 37, 9, 2557-2564.	$91,200 for 50 year-old $53,800 for 60 year-old	MEPS and NHIS based national number.	$119,919 for 50 year-old $107,712 for 60 year-old	We also use a MEPS and NHIS based national number.	Year 2012 dollars, 3% discount rate

Chapter 5: Procedures to Avoid Double Counting Benefits

We have found that many evaluations of programs and policies measure multiple outcomes. It is desirable, of course, to calculate benefits across multiple outcomes to draw a comprehensive conclusion about the total benefits of a program or policy. To do this, however, runs the risk of double counting certain outcome measures that really are alternative gauges of the same underlying effect.

For example, high school graduation and standardized test scores are two outcomes that may both be measured in a typical program evaluation. As described in Chapter 4, we have methods to monetize both of these outcomes individually; both lead to increased earnings in the labor market. These two outcomes, however, are likely to be, at least in part, measures of the same development in a person's human capital. If we simply add the separately calculated labor market benefits of each outcome, we would likely double count at least some of the same improved human capital generated by the program.

To avoid double counting program benefits, we have developed rules—we call them "trumping" rules—to reduce the chance that this will occur. This Chapter describes our procedures.

5.1 Trumping Rules 1, 2, and 3: When Multiple Outcome Measures Reflect the Same Underlying Construct

As noted, certain outcomes are likely to be alternative ways of measuring the same attribute. In the WSIPP benefit-cost model, we have identified the following outcomes that are alternative ways of measuring the same attribute.

- For human capital outcomes, we assume that the labor market gains from: 1) increases in student test scores; 2) increases in high school graduation; 3) increases in the number of years of education; 4) decreases in substance abuse; and 5) decreases in mental illness, reflect different measures of the same underlying construct that affect labor market performance.
- For health care outcomes, we assume that health care costs stemming from changes in: 1) high school graduation; 2) substance abuse; and 3) mental illness, reflect different measures of the same underlying construct that affect health care costs.

When a program has multiple outcomes *in either of these two categories*, we apply one of the following three trumping rules, depending on which situation applies:

- ✓ **Trumping Rule 1: The Biggest Winner.** When a topic has multiple favorable alternative outcomes and no undesirable (i.e., iatrogenic) outcomes, we determine the expected present value of benefits of each alternative outcome and select the one with the largest present value of benefits, and drop the other outcomes.[208] For example, if a program measures a gain in student test scores AND a gain in high school graduation rates, we compute the expected benefits from the present value of labor market earnings for both outcomes, and then select the outcome with the largest gain in present value benefits, and drop the other outcome from the benefit-cost analysis.
- ✓ **Trumping Rule 2: The Biggest Winner and the Biggest Loser.** When a topic has at least one favorable and at least one unfavorable (i.e., iatrogenic) alternative outcomes, we determine the expected present value of benefits of each favorable alternative outcome and the expected present value of losses of each unfavorable outcome and we then add together the outcome with the largest present value of benefits AND the outcome with the largest unfavorable outcome, ignoring the other outcomes. For example, if a program measures a gain in student test scores (a favorable outcome) and a reduction in high school graduation rates (an unfavorable outcome) we compute the expected present value of labor market gains from the test score outcome and the expected present value loss from the graduation outcome, and then we add these two together. Any other competing outcomes are dropped from the benefit-cost analysis.

[208] When determining which outcomes trump others, we implement these rules by running a single benefit-cost case where all inputs are taken at their modal values.

- ✓ **Trumping Rule 3: The Smallest Loser.** When a topic has multiple unfavorable alternative outcomes and no favorable outcomes, we determine the expected present value loss of each alternative outcome and select the one with the smallest present value loss. For example, if a program measures a reduction in student test scores (an unfavorable outcome) and a reduction in high school graduation rates (an unfavorable outcome) we compute the expected present value of labor market losses from both outcomes, and then we select the outcome with the smallest present value loss, and drop the other outcome from the benefit-cost analysis.

5.2 Trumping Rule 4: When There are Direct and Linked Paths to the Same Outcome

In addition to the possibility of double counting multiple outcome measures that reflect a common underlying construct, an additional trumping rule is needed when there are multiple paths to the same outcome.

As noted in this document, the WSIPP benefit-cost model monetizes changes to outcomes measured in one of two ways: a) directly from program evaluations that measure an outcome of interest, or b) from "linkage" studies that measure how a change in one outcome leads to a change in a second outcome.

Our fourth trumping rule is that in situations where the same outcome is measured from both direct and linked processes, we only monetize the outcome measured directly by the program evaluations. For example, a meta-analytic review of program evaluations may indicate that a home visiting program affects a) child abuse and neglect and b) high school graduation. Separately, our analysis of longitudinal linkage studies establishes that youth who are abused have a reduced probability of graduating from high school. In this example, we have two paths to the high school graduation outcome—the graduation outcome measured directly in the program evaluations and the graduation outcome measured in the linkage studies tracing the relationship be child abuse and graduation.

Our fourth trumping rule is this: in situations where are direct and linked paths to the same outcome, we only monetize the outcome directly measured in the program evaluation studies and we ignore the results of the measured linkage studies.

Chapter 6: Procedures to Estimate Program Costs

The WSIPP benefit-cost model implements a standard economic calculation of the expected worth of an investment by computing the net present value (NPV) of a stream of estimated benefits and costs that occur over time, as described with equation 6.1.

$$(6.1) \quad NPV_{tage} = \sum_{y=tage}^{N} \frac{Q_y \times P_y - C_y}{(1 + Dis)^y}$$

The procedures to produce, Qy—the outcomes achieved by the program or policy in a year y—were described in Chapters 2 and 3. The Py term—the price per unit of the outcome in year y—was discussed in Chapter 4. This Chapter describes the Cy term—the cost of producing the outcome 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.

Most of the program evaluations we review do not report information on the costs to implement a program. The focus of most program evaluations is on whether a program achieved outcomes, not on the costs of running a program.

For benefit-cost purposes, however, a program cost is needed.

To construct program cost estimates, we use several strategies and principles. These include the following:

- If the program evaluations we have meta-analyzed reflect a program currently in place in Washington, then we may collect program cost information from the relevant operating agency in Washington. We convert the program cost into a per participant number, usually an average cost, and use that cost estimate in our benefit-cost calculations.
- If the program evaluations we have meta-analyzed contain information on the number of "physical resource units" used by the program, then we summarize those units. For example, program evaluations of a K–12 tutoring program may report the number of sessions that a teacher works with a student, the number of hours per session, and the amount of preparation time for the teacher. We would use these physical unit parameters and then apply the average hourly cost of a teacher in Washington (information we obtain from other sources) to produce an estimate of the average cost of the tutoring program.
- Some programs or policy changes involve capital costs in addition to operating costs. When relevant, we include capital costs, expressed on an amortized per-participant basis.
- Depending on the design of particular program evaluations, we sometimes compare program participants to no-treatment comparison groups; in this case the comparison group would cost $0. In other evaluations, treated participants are compared to people who receive "treatment-as-usual." In this case, we use information from the program evaluations and/or Washington State data (as described in the first two points above) to estimate non-zero per-participant cost for the comparison group.
- Since our effect sizes are calculated on an intent-to-treat basis, it is important to construct the program cost parameters in a similar fashion. That is, the per-participant program costs represent the cost of the average person who enters the program, rather than the cost of a participant who completes the program.
- In addition to a per participant cost estimate, we also note the year in which the dollars are denominated.
- We also note the number of years over which the program costs are incurred, so that programs that involve multiple years of per-participant spending can be present valued with equation 6.1.

For each topic, the user enters seven pieces of information describing the program cost.

1. Treatment group: annual cost per participant.
2. Treatment group: the number of years over which the annual cost is incurred.
3. Treatment group: the year in which the cost estimate is denominated.
4. Comparison group: annual cost per participant.
5. Comparison group: the number of years over which the annual cost is incurred.
6. Comparison group: the year in which the cost estimate is denominated.
7. A percentage range, up and down, around the cost per participant estimates. The range is used in Monte Carlo simulation and is modeled with random draws from a triangular probability density distribution.

Chapter 7: Procedures to Access Risk and Uncertainty

Thus far in this Technical Documentation, we have focused on the single-point estimates of benefits and costs for 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 this single point estimate? If we vary the inputs, how often will costs exceed benefits, rather than the other way around?

WSIPP's benefit-cost model includes many inputs and assumptions, and there is significant risk and uncertainty around many of these factors. If the factors 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 thousands 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 compute an expected net present value, an expected benefit-cost ratio, and a straightforward measure of investment risk: for any program, what percentage of the time can we expect benefits to exceed costs? That is, our key measure of risk is this: after running the model 10,000 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)?

In Washington State, beginning in 2013, the legislature has directed WSIPP to create "inventories" of "evidence-based," "research-based" and "promising" programs and practices for several policy areas. We evaluate programs in each of these policy areas against the definitions; one criterion for meeting the "evidence-based" definition is that a program must "break-even," or, benefits must exceed costs, in at least 75% of the 10,000 Monte Carlo simulation runs. When the results fall between 73% and 77%, we re-run those programs 100,000 times to get a more precise estimate. We base this range on an analysis of ten programs that fell close to the 75% criterion; for each, we ran 100 independent 10,000 Monte Carlo simulation runs and recorded the "break-even" statistic for each. Calculating the minimum and maximum break-even for each program produced ranges between 1% and 3%, so we defined our range for re-running by adding and subtracting 2 percentage points from the 75% criterion.

7.1 Key Inputs Varied in the Monte Carlo Simulation Analysis

Potentially, all inputs to WSIPP'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 probability density distributions around the following list of inputs.

Program Effect Sizes. As described in Chapters 2 and 3, 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 its standard error by sampling from a normal probability density distribution.

Linked Effect Sizes. Chapters 2 and 3 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 size and its standard error to sample from a normal probability density distribution.

Discount Rates. Three different rates of discount (low, modal, and high) are used evaluate future benefits and costs in present value terms. In a single run of the model, the modal discount rate is used. In Monte Carlo 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 and include the following:

- Unit changes produced by program (and linked) effect sizes—normal distribution
- Program costs—triangular distribution
- Crime victimization costs—triangular distribution
- Criminal justice system costs—triangular distribution
- Crime police and prison elasticities—normal distribution
- Discount rate—triangular distribution
- Value of a statistical life—triangular distribution
- Deadweight cost of taxation—triangular distribution
- Labor market earnings from reduction in substance abuse/dependence, mental health disorders, CAN, and health conditions (alcohol disorders, regular tobacco smoking, cannabis disorders, and non-cannabis illicit drug disorders; depression, anxiety, SMI, and PTSD; CAN prevention and intervention; obesity and diabetes)—gamma/normal/lognormal distribution
- Expected higher education cost escalation—triangular distribution
- Expected health care cost escalation—triangular distribution
- Expected health care costs per instance of depression, anxiety, PTSD, serious mental illness, disruptive behavior, ADHD disorder, obesity, or diabetes—normal 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 for varying education levels—normal distribution
- Human capital economic externalities of education—triangular distribution

7.2 Computational Procedures to Carry Out the Simulation

Since the benefit-cost model is housed in Microsoft Excel® and uses spreadsheet formulas and 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 Rand function with the following procedure:

$$(7.1)\quad RandomDraw = RAND()$$

Next, the distribution for a particular probability distribution 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 the following procedure:

$$(7.2)\quad EffectSize_r = NORMINV(RandomDraw, EffectSizeMean, EffectSizeStandardError)$$

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 entered by the user.

$$(7.3)\quad If\ RandomDraw < \frac{(Mode - Min)}{(Max - Min)}\ then\ DISRATE = Min + \sqrt{RandomDraw \times (Mode - Min) \times (Max - Min)}$$

$$(7.4)\quad If\ RandomDraw \geq \frac{(Mode - Min)}{(Max - Min)}\ then\ DISRATE = Max - \sqrt{(1 - RandomDraw) \times (Max - Mode) \times (Max - Min)}$$

Chapter 8: The WSIPP Portfolio Tool

WSIPP constructed an analytical portfolio tool for the Washington Legislature to help identify evidence-based programming and policy options to improve outcomes for people in Washington State, as well as to reduce taxpayer and other societal costs. This portfolio tool is based on the sentencing tool developed by WSIPP in 2010[209] but has been expanded to include several new outcomes, not just those relevant to criminal justice.[210] The goal of the tool is to help users analyze the net effects of many kinds of evidence-based programs and policies, and examine the impact of user-defined combinations of programs and policies on net cash flows and caseloads. Specifically, the tool is designed to examine how changes in the mix of policy and programming strategies can affect, at the state level, the following: 1) the number victimizations from crime; 2) the number of prison beds needed; 3) the number of child abuse and neglect cases; 4) the number of out-of-home placements for children in child welfare; 5) the number of high school graduates; and 6) costs and benefits to society over time.

Evidence-Based Program Portfolio. The portfolio analysis tool imports the eligible saved results of Monte-Carlo simulation from the benefit-cost model. The user selects eligible programs to be analyzed in the portfolio tool. The user then either enters or uses the saved portfolio specific inputs for the selected programs as described below. This allows for the user to combine a unique set of programs and policies into a single portfolio.
The WSIPP portfolio tool implements a three-step computational process:

1. First, the user must use the benefit-cost model to create Monte Carlo results for each program to estimate the program's ability to affect outcomes and related taxpayer and societal savings;
2. Within the portfolio program, results of an overall portfolio of programming and policy resources are tallied (over a 50-year time frame); and
3. Sensitivity analysis is conducted by simulating uncertainty in the analysis using a Monte Carlo approach.

8.1 Estimating the Expected Benefits and Costs of Programs and Policies

Any program or policy in the WSIPP benefit-cost model can be run using a Monte Carlo approach. First, the mean, per-participant cash flows from the benefit-cost model are stored for each year in a participant's projected lifetime. The standard deviations from these means are also stored. Second, the mean per-participant "person counts" and their standard deviations are also stored for each year in a participant's projected lifetime. The person counts currently have five types: projected per-participant changes in prison average daily population, crime victimizations, high school graduates, child abuse and neglect cases, and out-of-home placements in child welfare. These counts underlie the benefit and cost calculations in the crime, child welfare, and high school graduation areas, detailed in Chapters 4.1, 4.2, 4.3, and 4.7.

Key parameters that are allowed to vary in the individual benefit-cost model are described in Chapter 7.1.

8.2 Preparing Programs and Policies for Portfolio Analysis

In addition to the results of a Monte Carlo simulation from the benefit-cost model, the portfolio analysis also requires several other pieces of information for each program or policy. Numbers for each policy are calculated on a per participant basis. The Portfolio tool requires the number of participants (slots) entering each program for each year that the program will be evaluated in the portfolio.

One important concept for long term portfolio analysis is that of diminishing returns. This is the precept that, as a program serves more and more of its eligible population (that is, as it reaches market saturation), the effectiveness of the program for each new participant may be reduced. The tool requires three pieces of information to model diminishing returns: the current annual funded participants in each program, the maximum number of annual eligible participants, and how effective the

[209] Aos, S., & Drake, E. (2010). *WSIPP's benefit-cost tool for states: Examining policy options in sentencing and corrections*. (Doc. No. 10-08-1201). Olympia: Washington State Institute for Public Policy.
[210] The high school graduation portion of the portfolio model was funded by the MacArthur Foundation, and the child welfare component was funded by the Pew Charitable Trusts.

program is expected to be at maximum capacity (the "diminishing returns factor," expressed as a decimal between zero and one where one means that there is as effective at the last eligible program participant as the first, while zero means the program is completely ineffective when it serves at the maximum level). The user is also able to estimate the variability expressed as percentage of the chosen diminishing returns factor; the variability is modeled with a triangular distribution in the portfolio Monte Carlo simulation.

Finally, the user is also required to enter an adjustment for each specific program, given what he or she knows about the mix of programs and policies in a given portfolio scenario. If the user had a portfolio which included several programs for high-to-moderate risk adult offenders (for example), the user might enter a lower or higher number to reflect the fact that individuals might receive more than one treatment and those treatments may not have fully independent effects on outcomes. A number less than one would indicate that if a participant participates in several programs, the combined effect will be less that the simple addition of the two individual program effects, while a number greater than one would indicate that the combined effects of multiple programs would be greater than the individual sum of each program's contributions.

8.3 Combining Results of a Portfolio of Programs and Policies

Using the previously stored results for the programs selected for the portfolio, the tool conducts a simple summation over time. For all programs in a portfolio, N, and for each follow-up year of investment i, the total change expected in a "person" outcome (e.g., prison beds, crime victimizations, child abuse and neglect cases, out-of-home placements) is the sum of the change in that person outcome for program p in investment year y, from follow up year one to i, multiplied by three factors: the number of slots funded in the follow up year for that program ($AddSlotspy$), the multiple-program adjustment factor for the program ($AdjFactorp$), and by the diminishing returns factor computed for that year ($DRFactorpy$).

$$(8.1) \quad \Delta Person_i = \sum_{p=1}^{N} \sum_{y=1}^{i} \Delta Person_{p(i-y+1)} * AddSlots_{py} * AdjFactor_p * DRFactor_{py}$$

We use Microsoft Excel's native future value (FV) and rate (RATE) functions to compute the diminishing returns multiplier ($DRFactory$) to adjust the expected effectiveness of a program, depending on how close the additional slots specified in the portfolio will bring us to maximum capacity. This factor may vary year to year, depending on the user-specified number of additional slots to be added.

DR is the expected level of effectiveness when the program reaches maximum capacity
$Current$ is the number of annual slots currently being funded statewide.
$AddSlots$ is the number of additional slots to be funded in year y.
$MaxCap$ is the maximum number of people in the state who meet the eligibility requirements for the program.

$$(8.2) \quad DRFactor_y = \frac{FV\left(RATE(99,0,-1,DR),\left(\frac{(Current + AddSlots_y)}{MaxCap} * 100,0,-1\right)\right) + FV\left(RATE(99,0,-1,DR),\left(\frac{Current}{MaxCap} * 100,0,-1\right)\right)}{2}$$

8.4 Portfolio Risk Analysis

Analyzing these program and policy investment scenarios involves a substantial amount of risk. While there is an increasingly strong evidentiary base of knowledge about what works to improve outcomes, there remains a considerable level of variation in particular estimates. To reflect this uncertainty, the third step in our portfolio modeling approach is designed to estimate the riskiness of any combination of policy options.

As with any investment decision, a risk-adverse 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 worse outcomes and economics, not better.

We estimate the known variability surrounding many of the inputs to the portfolio tool. Expected-value results of individual programs and policies are stored, using the variable parameters described in Chapter 7. We implement a Monte Carlo simulation approach in Excel, in which each time a scenario is run (the user selects the number of simulations to run), the tool draws randomly from the user-specified or model-generated probability distributions for the variables shown in the following table.

Exhibit 8.0

Parameters Allowed to Vary in Monte Carlo Simulation of a Portfolio Scenario

Portfolio-level parameter allowed to vary	Type of probability distribution
Portfolio-level variation	
Diminishing returns factor*	Triangular
Total annual cash flows	Normal
Change in crime victimizations	Normal
Change in prison ADP	Normal
Change in high school graduates	Normal
Change in child abuse and neglect cases	Normal
Change in child welfare out-of-home placements	Normal

* The specific parameters for this distribution are selected by the user.

The portfolio outputs are 50 years of total cash flows. In addition, we display expected values for changes in prison beds, crime victimizations, child abuse and neglect cases, out-of-home placements, and high school graduates.

Technical Documentation Appendix:

Estimates of Linked Relationships between Outcomes

As described earlier in this Technical Documentation, in addition to examining the impacts of a program on directly measured outcomes, we estimate the benefits of “linked” outcomes. For example, a program evaluation may measure the direct short-term effect of a child welfare program on child abuse outcomes but not the longer-term outcomes such as high school graduation. Other substantial bodies of research, however, have measured cause-and-effect relationships between being abused as a child and its effect on the odds of high school graduation. Using the same meta-analytic approach we describe in Chapter 2, we take advantage of this research and empirically estimate the causal “links” between two outcomes. In benefit-cost calculations, as described in Chapter 3, we then use these findings to project the degree to which a program is likely to have longer-term effects beyond those measured directly in program evaluations.

We list our current findings on these linkages in the three Exhibits in this Appendix: Exhibit 9.0 displays the meta-analytic results of each linkage we have estimated; Exhibit 9.1 shows the individual studies for each linkage; and Exhibit 9.3 is a list of citations for all of the studies in these meta-analyses of linked outcomes.

Exhibit 9.0

Linked Outcomes

Meta-Analytic Estimates of Standardized Mean Difference Effect Sizes

Estimated causal links between outcomes	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	p-value	Q-stat	p-value	ES	SE	p-value	ES	SE
ADHD, leading to...											
Grade retention	4	0.466	0.048	0.000	3.021	0.388	0.466	0.049	0.000	**0.466**	**0.049**
High school graduation	4	-0.311	0.025	0.000	0.108	0.991	-0.311	0.025	0.000	**-0.311**	**0.025**
Test scores-academic	see Externalizing composite										
Special education	see Externalizing composite										
Crime	see Externalizing composite										
Alcohol (problem use), leading to...											
High school graduation	7	-0.112	0.035	0.002	12.383	0.054	-0.166	0.061	0.007	**-0.166**	**0.061**
Employment	5	-0.337	0.023	0.000	88.448	0.000	-0.239	0.113	0.034	**-0.239**	**0.113**
Earnings given employment	1	-0.019	0.045	0.677	0.000	0.000	-0.019	0.045	0.677	**-0.019**	**0.045**
Crime	3	0.260	0.046	0.000	0.306	0.858	0.260	0.046	0.000	**0.260**	**0.046**
Alcohol Disorder, leading to...											
Earnings given employment	1	-0.051	0.040	0.204	0.000	0.000	-0.051	0.040	0.204	**-0.051**	**0.040**
Employment	2	-0.374	0.024	0.000	4.042	0.044	-0.360	0.051	0.000	**-0.360**	**0.051**
Crime	2	0.253	0.052	0.000	1.093	0.296	0.249	0.057	0.000	**0.249**	**0.057**
Alcohol Use < 18 years of age, leading to...											
High school graduation	6	-0.034	0.017	0.044	8.406	0.135	-0.039	0.030	0.201	**-0.039**	**0.030**
Crime	4	0.133	0.034	0.000	0.306	0.959	0.133	0.034	0.000	**0.133**	**0.034**
Anxiety Disorder, leading to...											
Employment	6	-0.165	0.029	0.000	27.176	0.000	-0.190	0.076	0.013	**-0.190**	**0.076**
Earnings given employment	2	-0.102	0.035	0.004	0.729	0.393	-0.102	0.035	0.004	**-0.102**	**0.035**
High school graduation	see Internalizing composite										
Grade retention	see Internalizing composite										
Births to < 18 Mother (child effect), leading to...											
Grade retention	3	0.229	0.039	0.000	0.939	0.625	0.229	0.039	0.000	**0.229**	**0.039**
High school graduation	3	-0.213	0.068	0.002	0.841	0.657	-0.213	0.068	0.002	**-0.213**	**0.068**
Tobacco (regular use)	1	0.052	0.137	0.706	0.000	0.000	0.052	0.137	0.706	**0.052**	**0.137**
Births to < 18 Mother (mother effect), leading to...											
Years of education completed	4	-0.073	0.030	0.016	0.178	0.981	-0.073	0.030	0.016	**-0.073**	**0.030**
High school graduation	4	-0.109	0.066	0.097	1.865	0.601	-0.109	0.066	0.097	**-0.109**	**0.066**
Public assistance	2	0.107	0.101	0.287	0.047	0.828	0.107	0.101	0.287	**0.107**	**0.101**
Child Abuse & Neglect, leading to...											
Depression	8	0.305	0.028	0.000	22.675	0.002	0.293	0.058	0.000	**0.293**	**0.058**
Alcohol (disordered use)	6	0.171	0.028	0.000	7.590	0.180	0.172	0.046	0.000	**0.172**	**0.046**
Illicit drugs (disordered use)	6	0.241	0.042	0.000	11.772	0.038	0.268	0.069	0.000	**0.268**	**0.069**
Test scores-academic	3	-0.270	0.062	0.000	2.278	0.320	-0.268	0.066	0.000	**-0.268**	**0.066**
Employment	3	-0.247	0.075	0.001	2.754	0.252	-0.258	0.094	0.006	**-0.258**	**0.094**
Years of education completed	1	-0.240	0.106	0.024	0.000	0.000	-0.240	0.106	0.024	**-0.240**	**0.106**
Anxiety (incl. OCD)	3	0.298	0.052	0.000	17.366	0.000	0.325	0.166	0.051	**0.325**	**0.166**
Obesity	5	0.022	0.018	0.242	9.052	0.060	0.042	0.039	0.283	**0.042**	**0.039**
PTSD	1	0.836	0.199	0.000	0.000	0.000	0.836	0.199	0.000	**0.836**	**0.199**
High school graduation	5	-0.412	0.048	0.000	14.308	0.006	-0.404	0.098	0.000	**-0.404**	**0.098**
Grade retention	1	0.446	0.102	0.000	0.000	0.000	0.446	0.102	0.000	**0.446**	**0.102**
Disruptive behavior	1	0.460	0.391	0.239	0.000	0.000	0.460	0.391	0.239	**0.460**	**0.391**
Special education	1	0.389	0.036	0.000	0.000	0.000	0.389	0.036	0.000	**0.389**	**0.036**
Tobacco (regular use)	1	0.387	0.123	0.002	0.000	0.000	0.387	0.123	0.002	**0.387**	**0.123**
Crime	11	0.532	0.034	0.000	35.330	0.000	0.542	0.071	0.000	**0.542**	**0.071**

Exhibit 9.0 (Continued)

Linked Outcomes

Meta-Analytic Estimates of Standardized Mean Difference Effect Sizes

Estimated causal links between outcomes	Number of effect sizes	Meta-analytic results before adjusting effect sizes: Fixed effects model, Weighted mean effect size & p-value			Fixed effects model, Homogeneity test (p-value to reject homogeneity)		Random effects, Weighted mean effect size & p-value			Adjusted effect size and standard error used in the benefit-cost analysis	
		ES	SE	p-value	Q-stat	p-value	ES	SE	p-value	ES	SE
Cannabis use <18 years of age, leading to...											
Crime	1	0.271	0.130	0.038	0.000	0.000	0.271	0.130	0.038	**0.271**	**0.130**
High school graduation	13	-0.180	0.016	0.000	109.830	0.000	-0.235	0.064	0.000	**-0.235**	**0.064**
Cesarean Section, leading to...											
Hospital readmissions	1	0.379	0.010	0.000	0.000	0.000	0.379	0.010	0.000	**0.379**	**0.010**
Crime (non-offender pop), leading to...											
High school graduation	6	-0.421	0.029	0.000	23.957	0.000	-0.505	0.079	0.000	**-0.505**	**0.079**
Crime (offender pop), leading to...											
High school graduation	4	-0.174	0.043	0.000	6.516	0.089	-0.191	0.066	0.004	**-0.191**	**0.066**
Depression, leading to...											
Employment	11	-0.296	0.018	0.000	97.244	0.000	-0.336	0.064	0.000	**-0.336**	**0.065**
Earnings given employment	3	-0.022	0.021	0.284	1.083	0.582	-0.022	0.021	0.284	**-0.022**	**0.021**
High school graduation	see Internalizing composite										
Grade retention	see Internalizing composite										
Diabetes, leading to...											
Employment	4	-0.220	0.013	0.000	18.479	0.000	-0.291	0.047	0.000	**-0.291**	**0.047**
Earnings given employment	3	-0.027	0.030	0.366	0.417	0.812	-0.027	0.030	0.366	**-0.027**	**0.030**
Nursing home	8	0.212	0.008	0.000	20.497	0.005	0.210	0.046	0.000	**0.210**	**0.046**
Disruptive Behavior Disorder, leading to...											
High school graduation	5	-0.429	0.029	0.000	6.063	0.195	-0.452	0.046	0.000	**-0.452**	**0.046**
Grade retention	4	0.273	0.055	0.000	1.155	0.764	0.273	0.055	0.000	**0.273**	**0.055**
Test scores-academic	see Externalizing composite										
Special education	see Externalizing composite										
Crime	see Externalizing composite										
Drug Disorder, leading to...											
Employment	5	-0.270	0.033	0.000	12.470	0.014	-0.293	0.059	0.000	**-0.293**	**0.059**
Crime	2	0.304	0.056	0.000	0.072	0.788	0.304	0.056	0.000	**0.304**	**0.056**
Externalizing behavior symptoms, leading to...											
High school graduation	3	-0.225	0.029	0.000	1.261	0.532	-0.225	0.029	0.000	**-0.225**	**0.029**
Externalizing Composite (includes conduct disorder & ADHD), leading to...											
Special education	2	0.398	0.091	0.000	0.047	0.828	0.398	0.091	0.000	**0.398**	**0.091**
Test scores-academic	5	-0.145	0.020	0.000	35.066	0.000	-0.185	0.076	0.015	**-0.185**	**0.076**
Crime	8	0.328	0.035	0.000	12.107	0.097	0.340	0.056	0.000	**0.340**	**0.056**
High School Graduation, leading to...											
Crime	7	-0.194	0.024	0.000	6.281	0.392	-0.194	0.025	0.000	**-0.194**	**0.025**
Internalizing Composite (includes depression & anxiety), leading to...											
High school graduation	7	-0.109	0.027	0.000	9.137	0.166	-0.117	0.037	0.002	**-0.117**	**0.037**
Grade retention	2	0.266	0.052	0.000	0.564	0.453	0.266	0.052	0.000	**0.266**	**0.052**
Obesity, leading to...											
Employment	2	-0.028	0.013	0.030	3.971	0.046	-0.074	0.065	0.252	**-0.074**	**0.065**
Earnings given employment	1	-0.028	0.023	0.223	0.000	0.000	-0.028	0.023	0.223	**-0.028**	**0.023**
Nursing home	3	0.177	0.030	0.000	0.840	0.657	0.177	0.030	0.000	**0.177**	**0.030**
PTSD, leading to...											
Employment	4	-0.391	0.022	0.000	26.311	0.000	-0.357	0.102	0.000	**-0.357**	**0.102**
Serious Mental Illness, leading to...											
Employment	4	-0.583	0.023	0.000	54.112	0.000	-0.491	0.112	0.000	**-0.491**	**0.112**
Earnings given employment	1	-0.103	0.044	0.019	0.000	0.000	-0.103	0.044	0.019	**-0.103**	**0.044**
Crime*	4	0.139	0.028	0.000	2.728	0.435	0.139	0.028	0.000	**0.139**	**0.028**

Exhibit 9.0 (Continued)

Linked Outcomes

Meta-Analytic Estimates of Standardized Mean Difference Effect Sizes

Estimated causal links between outcomes	Number of effect sizes	Meta-analytic results before adjusting effect sizes: Fixed effects model: Weighted mean effect size & p-value			Homogeneity test (p-value to reject homogeneity)		Random effects: Weighted mean effect size & p-value			Adjusted effect size and standard error used in the benefit-cost analysis	
		ES	SE	p-value	Q-stat	p-value	ES	SE	p-value	ES	SE
Serious Mental Illness with Subtance Abuse (comorbid), leading to...											
Crime*	4	0.689	0.033	0.000	10.380	0.016	0.663	0.091	0.000	**0.663**	**0.091**
Smoking regularly, leading to...											
Earnings given employment	4	-0.056	0.006	0.000	11.809	0.008	-0.054	0.020	0.008	**-0.054**	**0.020**
High school graduation	5	-0.394	0.016	0.000	14.536	0.006	-0.351	0.055	0.000	**-0.351**	**0.055**
Employment	4	-0.035	0.007	0.000	7.902	0.048	-0.045	0.013	0.000	**-0.045**	**0.013**
Youth binge drinking, leading to...											
Crime	2	0.274	0.073	0.000	0.248	0.618	0.274	0.073	0.000	**0.274**	**0.073**

*The weighted average of these two effect sizes is used in the model (71% of the ES depends on the SMI ES, 29% depends on the comorbid SMI/substance abuse ES).

Exhibit 9.1

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Unadjusted effect size	Number in test condition group	Number in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	ADHD	**Grade retention**						
605	Fletcher & Wolfe, 2008	0.453	261	2643	53.98	53.77	1.000	**0.453**
12756	Galera et al., 2009	0.597	163	1101	92.78	92.18	1.000	**0.597**
13157	Currie & Stabile, 2008	0.347	359	3232	99.28	98.58	1.000	**0.347**
13170	Currie & Stabile, 2008	0.468	582	5240	179.41	177.15	1.000	**0.468**
	ADHD	**High school graduation**						
610	Fletcher & Wolfe, 2008	-0.305	262	2645	93.84	93.84	1.000	**-0.305**
9848	Breslau et al., 2008	-0.322	486	5100	201.05	201.05	1.000	**-0.322**
12755	Galera et al., 2009	-0.369	71	643	22.60	22.60	1.000	**-0.369**
13712	Breslau et al., 2011	-0.309	2966	26696	1328.59	1328.59	1.000	**-0.309**
	ADHD	**Crime**			(see Externalizing composite)			
	ADHD	**Test scores-academic**			(see Externalizing composite)			
	ADHD	**Special education**			(see Externalizing composite)			
	Alcohol (problem use)	**High school graduation**						
10570	Renna, 2008	-0.033	654	1310	247.90	64.67	1.000	**-0.033**
11955	Chatterji et al., 2005	-0.349	105	1002	47.51	30.79	1.000	**-0.349**
18258	Hawkins et al., 2013	-0.028	86	5314	25.00	19.45	1.000	**-0.028**
7910	Dee & Evans, 2003	-0.222	1717	5749	70.27	38.97	1.000	**-0.222**
18383	Yan & Brocksen, 2013	-0.303	290	1523	52.73	32.90	1.000	**-0.303**
18461	Chatterji, 2006	-0.070	1695	5909	344.50	69.78	1.000	**-0.070**
18462	Hill et al., 2000	-0.799	242	566	9.80	8.81	1.000	**-0.799**
	Alcohol (problem use)	**Employment**						
10597	Saffer & Dave, 2005	-0.082	210	6790	125.93	14.64	1.000	**-0.082**
7119	Mullahy & Sindelar, 1996 Terza, 2002 Auld, 2002	-0.589	1448	13035	678.09	16.17	1.000	**-0.589**
7161	Feng et al., 2001	0.028	647	7475	245.54	15.52	1.000	**0.028**
7167	MacDonald & Shields, 2004	-0.217	664	5980	298.82	15.70	1.000	**-0.217**
13757	French et al., 2011	-0.312	1910	18459	534.72	16.07	1.000	**-0.312**
	Alcohol (problem use)	**Earnings given employment**						
30885	Bray, 2005 Keng & Huffman, 2010	-0.019	835	2139	483.68	NaN	1.000	**-0.019**
	Alcohol (problem use)	**Crime**						
17706	Popovici et al., 2012	0.251	756	8820	282.45	282.45	1.000	**0.251**
18554	Hill et al., 2000	0.156	242	566	16.30	16.30	1.000	**0.156**
18589	Viner & RM, January 01, 2007	0.285	874	4037	173.22	173.22	1.000	**0.285**
	Alcohol disorder	**Earnings given employment**						
9644	Jones & Richmond, 2006	-0.051	798	2848	622.90	NaN	1.000	**-0.051**
	Alcohol disorder	**Employment**						
7118	Mullahy & Sindelar, 1996	-0.406	2232	21573	1224.11	207.54	1.000	**-0.406**
30766	Sangchai, 2006 French et al., 2011	-0.303	1799	17399	551.19	171.95	1.000	**-0.303**
	Alcohol disorder	**Crime**						
17705	Popovici et al., 2012	0.280	756	8820	290.07	236.57	1.000	**0.280**
13742	Van Dorn et al., 2012	0.145	2946	31707	74.75	70.63	1.000	**0.145**
	Alcohol use < 18 years of age	**High school graduation**						
10569	Renna, 2008	-0.113	1082	882	280.69	181.31	1.000	**-0.113**
7081	Ellickson et al., 1998	0.057	3279	1111	260.06	172.47	1.000	**0.057**
7088	Bray et al., 2000	0.177	1144	248	36.59	34.15	1.000	**0.177**
7911	Dee & Evans, 2003	-0.250	1419	4330	57.95	52.06	1.000	**-0.250**
18460	Chatterji, 2006	-0.052	3029	4575	462.40	242.99	1.000	**-0.052**
13719	Breslau et al., 2011	-0.030	8543	21119	2380.85	421.44	1.000	**-0.030**
	Alcohol use < 18 years of age	**Crime**						
18527	Ellickson et al., January 01, 2003	0.131	2523	846	632.54	632.54	1.000	**0.131**
18552	Green et al., 2011	0.163	186	516	64.83	64.83	1.000	**0.163**
18588	Newcomb & MD, September 01, 1989	0.235	549	298	19.98	19.98	1.000	**0.235**
18633	Wells et al., December 01, 2004	0.118	729	224	171.06	171.06	1.000	**0.118**

Exhibit 9.1 (Continued)

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Unadjusted effect size	Number in test condition group	Number in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	Anxiety disorder	**Employment**						
9633	Cornwell et al., 2009	-0.101	1128	9513	139.94	29.31	1.000	**-0.101**
11798	Gibb et al., 2010	-0.151	143	808	38.96	19.00	1.000	**-0.151**
7185	Ettner et al., 1997	-0.087	562	4064	163.09	30.20	1.000	**-0.087**
12345	Chatterji et al., 2009	-0.117	1168	10645	561.79	34.77	1.000	**-0.117**
12349	Baldwin et al., 2007	-0.583	294	9675	136.10	29.13	1.000	**-0.583**
30947	Burnett-Zeigler et al., 2013 Cowell et al., 2009	-0.103	1220	26521	123.47	28.51	1.000	**-0.103**
	Anxiety disorder	**Earnings given employment**						
12352	Baldwin et al., 2007	-0.143	294	10530	285.94	285.94	1.000	**-0.143**
30906	Marcotte & Wilcox-Gők, 2003 Ettner et al., 1997	-0.080	657	3372	529.36	529.36	1.000	**-0.080**
	Births to < 18 mother (child effect)	**Grade retention**						
26598	Moore et al., 1997	0.245	77	199	24.21	24.21	1.000	**0.245**
24348	Levine et al., 2007	0.326	420	385	84.82	84.82	1.000	**0.326**
7038	Angrist & Lavy, 1996	0.213	557	17238	539.16	539.16	1.000	**0.213**
	Births to < 18 mother (child effect)	**High school graduation**						
26612	Francesconi, 2008	-0.314	85	1098	53.59	53.59	1.000	**-0.314**
12783	Hoffman & Scher, 2008	-0.205	644	337	86.84	86.84	1.000	**-0.205**
12794	Manlove et al., 2008	-0.150	221	461	73.75	73.75	1.000	**-0.150**
	Births to < 18 mother (child effect)	**Tobacco (regular use)**						
12801	Francesconi, 2008	0.052	85	1098	53.17	0.00	1.000	**0.052**
	Births to < 18 mother (mother effect)	**Years of education completed**						
26727	Kane et al., 2013	-0.079	890	6980	789.11	789.11	1.000	**-0.079**
27228	Ashcraft et al., 2013	-0.059	1313	186	162.89	162.89	1.000	**-0.059**
27270	Webbink et al., 2009	-0.006	49	49	24.50	24.50	1.000	**-0.006**
27303	Fletcher & Wolfe, 2009	-0.065	561	148	117.14	117.14	1.000	**-0.065**
	Births to < 18 mother (mother effect)	**High school graduation**						
27225	Ashcraft et al., 2013	-0.037	1313	186	109.77	109.77	1.000	**-0.037**
11799	Fletcher & Wolfe, 2009	-0.241	563	148	71.15	71.15	1.000	**-0.241**
12785	Webbink et al., 2009	-0.065	77	77	25.41	25.41	1.000	**-0.065**
12800	Hoffman, 2008	-0.096	453	41	25.28	25.28	1.000	**-0.096**
	Births to < 18 mother (mother effect)	**Public assistance**						
11800	Fletcher & Wolfe, 2009	0.137	564	149	35.01	35.01	1.000	**0.137**
12799	Hoffman, 2008	0.091	762	69	63.25	63.25	1.000	**0.091**
	Child abuse & neglect	**Depression**						
10544	Scott et al., 2010	0.525	221	1923	68.96	32.19	1.000	**0.525**
21916	Chapman et al., 2004	0.411	2373	7087	511.72	54.00	1.000	**0.411**
22116	Fletcher, 2009	0.297	182	3840	80.39	34.48	1.000	**0.297**
22260	Brown et al., 1999	0.665	81	558	18.77	14.32	1.000	**0.665**
6850	Widom et al., 2007	0.145	676	520	139.30	42.12	1.000	**0.145**
12409	Thornberry et al., 2010	0.158	170	645	134.26	41.65	1.000	**0.158**
12433	Springer et al., 2007	0.156	234	1817	207.05	46.74	1.000	**0.156**
12946	Fergusson et al., 2008	0.266	162	839	76.54	33.75	1.000	**0.266**
	Child abuse & neglect	**Alcohol (disordered use)**						
10552	Scott et al., 2010	0.332	221	1923	55.11	44.87	1.000	**0.332**
22106	Horwitz et al., 2001	-0.058	315	271	88.99	65.01	1.000	**-0.058**
22107	Horwitz et al., 2001	0.214	322	239	85.85	63.32	1.000	**0.214**
6768	Fergusson & Lynskey, 1997	0.409	118	111	23.93	21.77	1.000	**0.409**
12408	Thornberry et al., 2010	0.171	170	645	134.21	86.24	1.000	**0.171**
12421	Shin et al., 2009	0.173	6729	6019	851.91	188.02	1.000	**0.173**

Exhibit 9.1 (Continued)

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Unadjusted effect size	Number in test condition group	Number in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	Child abuse & neglect	**Illicit drugs (disordered use)**						
10553	Scott et al., 2010	0.695	221	1923	42.61	25.75	1.000	**0.695**
22219	Huang et al., 2011	0.240	1279	603	118.98	42.05	1.000	**0.240**
22339	Arteaga et al., 2010	0.273	117	1091	43.58	26.09	1.000	**0.273**
6740	McGloin & Widom, 2001	0.135	676	520	193.43	48.67	1.000	**0.135**
12407	Thornberry et al., 2010	0.275	170	645	133.71	43.76	1.000	**0.275**
12949	Fergusson et al., 2008	0.113	162	839	38.87	24.33	1.000	**0.113**
	Child abuse & neglect	**Test scores-academic**						
10750	Topitzes et al., 2010	-0.220	135	990	118.50	98.72	1.000	**-0.220**
6760	Eckenrode et al., 1993	-0.383	206	206	101.15	86.38	1.000	**-0.383**
6786	Lansford et al., 2002	-0.145	50	387	44.23	41.15	1.000	**-0.145**
	Child abuse & neglect	**Employment**						
21862	Currie & Widom, 2010	-0.470	174	174	42.22	31.85	1.000	**-0.470**
22263	Covey et al., 2013	-0.156	124	1169	27.58	22.75	1.000	**-0.156**
23093	Mersky & Topitzes, 2010	-0.184	184	1178	107.54	58.79	1.000	**-0.184**
	Child abuse & neglect	**Years of education completed**						
21863	Currie & Widom, 2010	-0.240	179	179	88.86	88.86	1.000	**-0.240**
	Child abuse & neglect	**Anxiety (incl. OCD)**						
21980	Springer et al., 2007	0.165	234	1817	207.02	12.92	1.000	**0.165**
22232	Scott et al., 2010	0.649	221	1923	101.78	12.13	1.000	**0.649**
12947	Fergusson et al., 2008	0.157	162	839	57.23	11.10	1.000	**0.157**
	Child abuse & neglect	**Obesity**						
22136	Power et al., 2015	0.100	766	3373	322.37	147.43	1.000	**0.100**
22145	Power et al., 2015	-0.011	869	3270	328.68	148.73	1.000	**-0.011**
22199	Bentley & Widom, 2009	0.004	410	303	174.24	106.15	1.000	**0.004**
22220	Shin & Miller, 2012	0.010	4406	4066	2114.58	240.74	1.000	**0.010**
12420	Noll et al., 2007	0.543	84	89	23.51	21.64	1.000	**0.543**
	Child abuse & neglect	**PTSD**						
22262	Shenk et al., 2014	0.836	51	59	25.17	25.17	1.000	**0.836**
	Child abuse & neglect	**High school graduation**						
6729	Thornberry et al., 2001	-0.176	134	604	45.93	18.15	1.000	**-0.176**
6738	McGloin & Widom, 2001	-0.479	676	520	185.50	25.84	1.000	**-0.479**
6822	Lansford et al., 2007	-0.854	69	505	34.48	16.05	1.000	**-0.854**
6871	Boden et al., 2007	-0.158	171	800	64.46	20.48	1.000	**-0.158**
12770	Mersky & Topitzes, 2010	-0.407	179	1148	99.45	23.06	1.000	**-0.407**
	Child abuse & neglect	**Grade retention**						
6762	Eckenrode et al., 1993	0.446	379	394	96.71	NaN	1.000	**0.446**
	Child abuse & neglect	**Disruptive behavior**						
6769	Fergusson & Lynskey, 1997	0.460	118	111	6.55	0.00	1.000	**0.460**
	Child abuse & neglect	**Special education**						
7488	Jonson-Reid et al., 2004	0.389	3987	3953	767.88	NaN	1.000	**0.389**
	Child abuse & neglect	**Tobacco (regular use)**						
12774	Mersky & Topitzes, 2010	0.387	143	919	65.62	NaN	1.000	**0.387**

Exhibit 9.1 (Continued)
Linked Outcomes
Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation		Unadjusted effect size	Number in test condition group	Number in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	Child abuse & neglect	**Crime**							
21923	Cohen et al., 2004		0.530	51	579	30.48	14.91	1.000	**0.530**
21998	Currie & Tekin, 2012		0.414	512	1704	147.39	24.36	1.000	**0.414**
22109	Lansford et al., 2007		0.664	69	505	18.32	11.26	1.000	**0.664**
22181	Kazemian, et al., 2011		0.477	202	50	14.51	9.69	1.000	**0.477**
22265	Allwood & Widom, 2013		0.389	676	520	167.29	24.85	1.000	**0.389**
6718	English et al., 2002		0.600	877	877	235.64	25.97	1.000	**0.600**
6749	Stouthamer-Loeber et al., 2001		0.379	52	104	23.09	12.89	1.000	**0.379**
6860	Mersky & Reynolds, 2007		0.451	129	1275	49.30	18.33	1.000	**0.451**
7388	Lemmon, 1999		1.083	267	365	79.82	21.37	1.000	**1.083**
9057	Stouthamer-Loeber et al., 2002		0.635	83	179	31.62	15.18	1.000	**0.635**
12406	Thornberry et al., 2010		0.342	170	645	82.55	21.56	1.000	**0.342**
	Cannabis use <18 years of age	**High school graduation**							
18151	Green et al., 2010		-0.595	185	517	31.91	14.68	1.000	**-0.595**
18256	Hawkins et al., 2013		0.104	124	5276	34.25	15.16	1.000	**0.104**
7079	Ellickson et al., 1998		-0.074	860	3530	197.85	23.91	1.000	**-0.074**
7086	Bray et al., 2000		-0.508	677	715	26.31	13.37	1.000	**-0.508**
7151	Fergusson & Horwood, 1997		-0.385	180	755	73.90	19.88	1.000	**-0.385**
7159	Brook et al., 2002		-0.217	100	1048	27.21	13.60	1.000	**-0.217**
7389	McCaffrey et al., 2009		-0.112	276	2482	20.39	11.65	1.000	**-0.112**
18320	Legleye et al., 2010		0.078	13026	16367	765.41	26.27	1.000	**0.078**
12749	Yamada et al., 1996		-0.179	75	597	12.92	8.76	1.000	**-0.179**
12804	van Ours & Williams, 2009		-0.198	5931	5862	1992.44	26.83	1.000	**-0.198**
12811	Horwood et al., 2010		-0.480	1418	2176	337.61	25.17	1.000	**-0.480**
12815	Horwood et al., 2010		-0.162	407	1036	106.56	21.67	1.000	**-0.162**
12817	Horwood et al., 2010		-0.387	994	2176	267.06	24.68	1.000	**-0.387**
	Cannabis use <18 years of age	**Crime**							
18155	Green et al., 2010		0.271	185	517	58.90	NaN	1.000	**0.271**
	Cesarean section	**Hospital readmissions**							
26749	Liu et al., 2002		0.379	483263	2169463	10494.91	10494.91	1.000	**0.379**
	Crime (non-offender pop)	**High school graduation**							
28344	Hjalmarsson, 2008		-0.304	1222	6195	634.27	36.69	1.000	**-0.304**
12777	Tanner et al., 1999		-0.403	478	1882	130.90	30.01	1.000	**-0.403**
12823	Hirschfield, 2009		-0.666	216	2039	31.60	17.45	1.000	**-0.666**
12930	Apel & Sweeten, 2009		-0.623	400	4649	233.53	33.38	1.000	**-0.623**
13721	Webbink et al., 2012		-0.595	224	2028	99.94	28.02	1.000	**-0.595**
13722	Kirk & Sampson, 2009		-0.576	79	115	27.52	16.12	1.000	**-0.576**
	Crime (offender pop)	**High school graduation**							
28345	Hjalmarsson, 2008		-0.250	169	296	68.76	42.23	1.000	**-0.250**
28350	Apel & Sweeten, 2009		-0.079	656	1036	199.49	70.68	1.000	**-0.079**
28351	Apel & Sweeten, 2009		-0.360	315	508	113.66	55.76	1.000	**-0.360**
30202	Hjalmarsson, 2008		-0.127	465	466	154.48	64.06	1.000	**-0.127**

Exhibit 9.1 (Continued)

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Unadjusted effect size	Number in test condition group	Number in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	Depression	**Employment**						
9632	Cornwell et al., 2009	-0.401	724	9917	60.49	18.91	1.000	**-0.401**
11797	Gibb et al., 2010	-0.351	143	808	46.62	17.30	1.000	**-0.351**
17799	Peng et al., 2013	-0.081	1386	13841	681.70	26.44	1.000	**-0.081**
7184	Ettner et al., 1997	-0.328	454	4172	170.63	23.69	1.000	**-0.328**
7193	Farahati et al., 2003	-0.255	74	438	32.84	14.97	1.000	**-0.255**
7202	Savoca & Rosenheck, 2000	-0.315	79	1338	31.66	14.72	1.000	**-0.315**
7216	Alexandre & French, 2001	-0.527	384	890	144.46	23.11	1.000	**-0.527**
12344	Chatterji et al., 2009	-0.310	1709	10104	861.65	26.66	1.000	**-0.310**
12346	Tian et al., 2005	-0.150	459	5239	280.00	25.05	1.000	**-0.150**
12348	Baldwin et al., 2007	-0.690	703	10121	328.19	25.38	1.000	**-0.690**
30948	Burnett-Zeigler et al., 2013 Cowell et al., 2009	-0.290	942	26799	285.73	25.09	1.000	**-0.290**
	Depression	**Earnings given employment**						
17800	Peng et al., 2013	-0.004	1386	13841	1259.51	1259.51	1.000	**-0.004**
12351	Baldwin et al., 2007	-0.054	703	10121	657.28	657.28	1.000	**-0.054**
30905	Marcotte & Wilcox-Gők, 2003 Ettner et al., 1997	-0.027	468	3560	411.63	411.63	1.000	**-0.027**
	Depression	**High school graduation**			(see Internalizing composite)			
	Depression	**Grade retention**			(see Internalizing composite)			
	Diabetes	**Employment**						
21552	Stewart et al., 2007	-0.458	1033	18042	167.21	79.65	1.000	**-0.458**
21577	Minor, 2013	-0.194	7808	95780	4453.54	147.07	1.000	**-0.194**
31015	Kahn, 1998 Ng et al., 2001	-0.256	1852	77416	970.80	131.49	1.000	**-0.256**
31016	Kahn, 1998 Tunceli et al., 2005	-0.345	725	7652	306.18	101.62	1.000	**-0.345**
	Diabetes	**Earnings given employment**						
21578	Minor, 2013	-0.065	221	32302	219.41	219.41	1.000	**-0.065**
21584	Songer et al., 1989	0.000	127	127	63.50	63.50	1.000	**0.000**
21591	Kahn, 1998	-0.019	959	8738	864.14	864.14	1.000	**-0.019**
	Diabetes	**Nursing home**						
21596	Valiyeva et al., 2006	0.714	173	3353	47.77	33.92	1.000	**0.714**
21601	Valiyeva et al., 2006	0.246	255	2681	89.90	50.84	1.000	**0.246**
21603	Harris & Cooper, 2006	0.213	21148	116484	15482.00	116.12	1.000	**0.213**
21607	Braunseis et al., 2011	0.079	460	1841	54.08	36.98	1.000	**0.079**
21608	Luppa et al., 2010	-0.234	151	603	14.95	13.26	1.000	**-0.234**
21609	Stineman et al., 2012	0.285	928	6908	157.81	67.19	1.000	**0.285**
21610	Andel et al., 2007	0.135	486	1457	239.50	78.60	1.000	**0.135**
21634	Banaszak-Holl et al., 2004	0.110	901	5775	210.22	75.17	1.000	**0.110**
	Disruptive behavior disorder	**High school graduation**						
9847	Breslau et al., 2008	-0.555	380	5206	191.32	114.64	1.000	**-0.555**
8027	Fergusson & Lynskey, 1998	-0.333	83	886	41.02	35.88	1.000	**-0.333**
12757	Galera et al., 2009	-0.438	71	643	20.95	19.52	1.000	**-0.438**
13710	Breslau et al., 2011	-0.386	1513	28149	767.78	208.41	1.000	**-0.386**
13760	Porche et al., 2011	-0.525	287	2245	129.58	89.18	1.000	**-0.525**
	Disruptive behavior disorder	**Grade retention**						
12758	Galera et al., 2009	0.292	163	1101	95.12	95.12	1.000	**0.292**
13148	Currie & Stabile, 2008	0.386	183	3403	57.17	57.17	1.000	**0.386**
13169	Currie & Stabile, 2008	0.235	297	5519	77.44	77.44	1.000	**0.235**
13755	Webbink et al., 2011	0.221	249	1971	101.73	101.73	1.000	**0.221**

Exhibit 9.1 (Continued)

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Unadjusted effect size	Number in test condition group	Number in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	Disruptive behavior disorder	**Crime**			(see Externalizing composite)			
	Disruptive behavior disorder	**Test scores-academic**			(see Externalizing composite)			
	Disruptive behavior disorder	**Special education**			(see Externalizing composite)			
	Drug disorder	**Employment**						
5574	Buchmueller & Zuvekas, 1998	-0.220	449	1651	178.80	57.81	1.000	**-0.220**
7105	Zuvekas et al., 2005	-0.171	929	8089	226.91	62.06	1.000	**-0.171**
7169	Alexandre & French, 2004	-0.285	926	553	226.15	62.01	1.000	**-0.285**
7171	French et al., 2001	-0.271	379	9242	215.99	61.22	1.000	**-0.271**
7190	Ettner et al., 1997	-0.624	148	4478	78.02	40.78	1.000	**-0.624**
	Drug disorder	**Crime**						
17708	Popovici et al., 2012	0.308	756	8820	297.36	297.36	1.000	**0.308**
13743	Van Dorn et al., 2012	0.250	693	33960	22.58	22.58	1.000	**0.250**
	Externalizing composite (includes conduct disorder & ADHD)	**Special education**						
606	Fletcher & Wolfe, 2008	0.380	231	2339	65.31	65.31	1.000	**0.380**
28630	Currie & Stabile, 2008	0.420	206	2527	54.48	54.48	1.000	**0.420**
	Externalizing composite (includes conduct disorder & ADHD)	**Test scores-academic**						
27215	Massetti et al., 2008	-0.143	85	130	51.17	23.42	1.000	**-0.143**
27237	Turney & McLanahan, 2015	0.044	821	1481	527.11	39.92	1.000	**0.044**
28603	Currie & Stabile, 2008	-0.449	180	2199	148.83	33.48	1.000	**-0.449**
28604	Currie & Stabile, 2008	-0.245	194	2380	161.74	34.09	1.000	**-0.245**
12009	Currie et al., 2010	-0.168	1739	48665	1678.16	42.11	1.000	**-0.168**
	Externalizing composite (includes conduct disorder & ADHD)	**Crime**						
28580	Currie & Stabile, 2008	0.132	243	2980	137.64	60.78	1.000	**0.132**
28581	Currie & Stabile, 2008	0.234	172	2106	96.37	51.11	1.000	**0.234**
9481	Satterfield et al., 2007	0.535	169	64	25.45	20.63	1.000	**0.535**
12732	Fletcher & Wolfe, 2009	0.419	85	858	44.08	31.37	1.000	**0.419**
12735	Copeland et al., 2007	0.339	125	1296	44.89	31.78	1.000	**0.339**
12742	Fergusson et al., 2005	0.763	46	927	17.42	15.02	1.000	**0.763**
12822	Murray et al., 2010	0.360	1090	7296	427.36	86.74	1.000	**0.360**
13754	Webbink et al., 2011	0.501	98	778	26.65	21.41	1.000	**0.501**
	Externalizing, clinical	**High school graduation**						
612	McLeod & Kaiser, 2004	-0.273	57	367	22.04	22.04	1.000	**-0.273**
12008	Currie et al., 2010	-0.234	1739	48665	1090.49	1090.49	1.000	**-0.234**
13753	Webbink et al., 2011	-0.124	248	1970	110.03	110.03	1.000	**-0.124**
	High school graduation	**Crime**						
7067	Lochner & Moretti, 2004	-0.183	102	102	50.79	50.24	1.000	**-0.183**
7069	Lochner & Moretti, 2004	-0.146	2162	540	431.59	395.21	1.000	**-0.146**
12775	Ou & Reynolds, 2010	-0.211	374	359	119.63	116.65	1.000	**-0.211**
12795	Machin et al., 2011	-0.212	85	85	42.01	41.64	1.000	**-0.212**
13720	Webbink et al., 2012	-0.147	1568	684	108.97	106.49	1.000	**-0.147**
13724	Bjerk, 2011	-0.293	1286	672	437.13	399.85	1.000	**-0.293**
13740	Van Dorn et al., 2012	-0.158	28987	5666	527.98	474.54	1.000	**-0.158**
	Internalizing composite (includes depression & anxiety)	**High school graduation**						
611	McLeod & Kaiser, 2004	-0.210	75	349	25.99	24.08	1.000	**-0.210**
1841	Duchesne et al., 2008	-0.215	177	1640	93.56	72.79	1.000	**-0.215**
27128	Breslau et al., 2011	-0.048	3502	26161	602.93	212.33	1.000	**-0.048**
27133	Breslau et al., 2008	-0.211	1218	4368	326.92	163.67	1.000	**-0.211**
28578	Composite of Needham, 2009 and Fletcher, 2010	-0.144	876	7404	204.27	125.84	1.000	**-0.144**
12937	Fergusson & Woodward, 2002	-0.058	124	840	43.43	38.35	1.000	**-0.058**
13762	Porche et al., 2011	0.018	368	2164	111.40	83.14	1.000	**0.018**

Exhibit 9.1 (Continued)

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Unadjusted effect size	Number in test condition group	Number in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	Internalizing composite (includes depression & anxiety)	**Grade retention**						
13158	Currie & Stabile, 2008	0.313	640	2958	150.64	150.64	1.000	**0.313**
13171	Currie & Stabile, 2008	0.234	1038	4793	224.20	224.20	1.000	**0.234**
	Obesity	**Employment**						
21813	Han et al., 2009	-0.023	16305	95924	5624.17	146.32	1.000	**-0.023**
21825	Tunceli et al., 2006	-0.156	526	2419	232.35	91.24	1.000	**-0.156**
	Obesity	**Earnings given employment**						
31012	Baum et al., 2006 Dastan, 2011	-0.028	2750	5915	1831.78	0.00	1.000	**-0.028**
	Obesity	**Nursing home**						
21856	Elkins et al., 2006	0.159	917	4367	776.94	776.94	1.000	**0.159**
21859	Valiyeva et al., 2006	0.196	645	2881	103.02	103.02	1.000	**0.196**
31119	Valiyeva et al., 2006	0.224	537	2399	246.28	246.28	1.000	**0.224**
	PTSD	**Employment**						
17985	WSIPP, 2013	-0.440	2496	32157	1554.16	29.82	1.000	**-0.440**
18013	Resnick & SG, 2008	-0.128	925	4901	317.87	27.75	1.000	**-0.128**
18051	McCarren et al., 1995	-0.432	273	273	46.57	18.40	1.000	**-0.432**
30945	Zatzick et al., 1997 Savoca & Rosenheck, 2000	-0.479	278	1025	63.13	20.52	1.000	**-0.479**
	Serious mental illness	**Employment**						
17731	Burnett-Zeigler et al., 2013	-0.209	509	21898	195.92	19.21	1.000	**-0.209**
7189	Ettner et al., 1997	-0.331	479	4147	179.20	19.03	1.000	**-0.331**
12353	Baldwin et al., 2007	-0.770	804	9675	384.61	20.18	1.000	**-0.770**
13185	WSIPP, 2014	-0.626	1728	36141	1086.61	20.89	1.000	**-0.626**
	Serious mental illness	**Earnings given employment**						
12354	Baldwin et al., 2007	-0.103	542	9675	513.11	513.11	1.000	**-0.103**
	Serious mental illness	**Crime**						
11227	Fazel et al., 2010	0.091	2948	29479	286.00	286.00	1.000	**0.091**
11644	Fazel et al., 2009	0.140	6035	60333	916.37	916.37	1.000	**0.140**
7267	Steadman et al., 1998	0.167	212	428	14.52	14.52	1.000	**0.167**
13745	Van Dorn et al., 2012	0.300	2531	24951	79.09	79.09	1.000	**0.300**
	Serious mental illness with subtance abuse (comorbid)	**Crime**						
11228	Fazel et al., 2010	0.845	795	7950	229.53	43.87	1.000	**0.845**
11645	Fazel et al., 2009	0.650	1960	19656	655.57	50.09	1.000	**0.650**
7268	Steadman et al., 1998	0.165	54	91	10.27	8.64	1.000	**0.165**
13746	Van Dorn et al., 2012	0.502	592	24951	28.24	18.57	1.000	**0.502**
	Smoking regularly	**Earnings given employment**						
9658	Anger & Kvasnicka, 2010	-0.164	819	1149	476.61	315.67	1.000	**-0.164**
12808	Jofre-Bonet et al., 2005	-0.061	31105	88778	23026.42	898.31	1.000	**-0.061**
12931	Braakmann, 2008	-0.013	3611	8647	2547.18	683.82	1.000	**-0.013**
30943	Baum et al., 2006 Cowan & Schwab, (2011). Dastan, 2011	-0.030	2468	5355	1660.29	598.06	1.000	**-0.030**

Exhibit 9.1 (Continued)

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Unadjusted effect size	Number in test condition group	Number in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	Smoking regularly	**High school graduation**						
18257	Hawkins et al., 2013	-0.455	605	4795	241.17	74.58	1.000	**-0.455**
7080	Ellickson et al., 1998	-0.191	2182	2208	297.05	79.19	1.000	**-0.191**
7087	Bray et al., 2000	-0.345	926	466	30.05	23.51	1.000	**-0.345**
18385	Yan & Brocksen, 2013	-0.321	490	1323	74.57	44.10	1.000	**-0.321**
13718	Breslau et al., 2011	-0.411	9818	19844	3177.37	104.42	1.000	**-0.411**
	Smoking regularly	**Employment**						
12807	Jofre-Bonet et al., 2005	-0.020	31105	88778	12120.63	2183.12	1.000	**-0.020**
12819	Dastan, 2011	-0.073	4005	8004	1330.12	887.02	1.000	**-0.073**
13188	WSIPP, 2014	-0.047	11082	26357	4849.32	1718.89	1.000	**-0.047**
13189	WSIPP, 2014	-0.062	9064	22850	3234.45	1460.43	1.000	**-0.062**
	Youth binge drinking	**Crime**						
18554	Hill et al., 2000	0.156	242	566	16.30	16.30	1.000	**0.156**
18589	Viner & RM, January 01, 2007	0.285	874	4037	173.22	173.22	1.000	**0.285**

Merged cells have had the effect sizes of different studies combined in a separate fixed-effects meta-analysis

Exhibit 9.2

Citations used in Linked Outcomes from Exhibits 9.0 and 9.1

Alexandre, P., & French, 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.

Allwood, M.A., & Widom, C.S. (2013). Child abuse and neglect, developmental role attainment, and adult arrests. *Journal of Research in Crime and Delinquency, 50*(4), 551-578.

Andel, R., Hyer, K., & Slack, A. (2007). Risk factors for nursing home placement in older adults with and without dementia. *Journal of Aging and Health, 19*(2), 213-228.

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). *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). *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.

Ashcraft, A., Fernández-Val, I., & Lang, K. (2013). The consequences of teenage childbearing: Consistent estimates when abortion makes miscarriage non-random. *The Economic Journal, 123*(571), 875-905.

Auld, M.C. (2002). *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 <ahref="http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.7.6712">http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.7.6712

Baldwin, M.L., & Marcus, S.C. (2007). Labor market outcomes of persons with mental disorders. *Industrial Relations, 46*(3), 481-510.

Banaszak-Holl, J., Fendrick, A.M., Foster, N.L., Herzog, A.R., Kabeto, M.U., Kent, D.M., Straus, W.L., ... Langa, K.M. (2004). Predicting nursing home admission: estimates from a 7-year follow-up of a nationally representative sample of older Americans. *Alzheimer Disease and Associated Disorders, 18*, 2.

Baum, C.L., Ford, W.F., & Hopper, J.D. (2006). The obese smoker's wage penalty. *Social Science Quarterly, 87*(4), 863-881.

Bentley, T., & Widom, C. S. (2009). A 30-year follow-up of the effects of child abuse and neglect on obesity in adulthood. *Obesity, 17*(10), 1900-1905.

Bjerk, D. (2011). *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 <ahref="http://ftp.iza.org/dp5995.pdf">http://ftp.iza.org/dp5995.pdf

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). *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.

Braunseis, F., Deutsch, T., Frese, T., & Sandholzer, H. (2011). The risk for nursing home admission (NHA) did not change in ten years—A prospective cohort study with five-year follow-up. *Archives of Gerontology and Geriatrics.*

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.

Brown, J., Cohen, P., Johnson, J. G., & Smailes, E. M. (December 01, 1999). Childhood Abuse and Neglect: Specificity of Effects on Adolescent and Young Adult Depression and Suicidality. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 12.

Buchmueller, T.C., & Zuvekas, S.H. Drug Use, Drug abuse, and labour market outcomes. *Health Economics, 7*, 229-45.

Burnett-Zeigler, I., Ilgen, M.A., Bohnert, K., Miller, E., Islam, K., & Zivin, K. (2013). The impact of psychiatric disorders on employment: results from a national survey (NESARC). *Community Mental Health Journal, 49*(3), 303-10.

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. (2006). Does alcohol use during high school affect educational attainment? Evidence from the National Education Longitudinal Study. *Economics of Education Review, 25*(5), 482-497.

Chatterji, P., & DeSimone, J. (2005). *Adolescent drinking and high school dropout* (Working Paper No. 11337). Cambridge: National Bureau of Economic Research.

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.

Cohen, P., Smailes, E., & Brown, J. (2004). *Effects of childhood maltreatment on adult arrests in a general population sample* (Document No. NCJ 199707). Washington, DC: National Institute of Justice.

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.

Covey, H.C., Menard, S., & Franzese, R.J. (2013). Effects of adolescent physical abuse, exposure to neighborhood violence, and witnessing parental violence on adult socioeconomic status. *Child Maltreatment, 18*(2), 85-97.

Cowan, B., & Schwab, B. (2011). The incidence of the healthcare costs of smoking. *Journal of Health Economics, 30*(5), 1094-1102.

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.

Currie, J., & Stabile, M. 2008. *Mental health in childhood and human capital* (Working Paper 13217). Cambridge, MA: National Bureau of Economic Research.

Currie, J., & Tekin, E. (2012). Understanding the Cycle: Childhood Maltreatment and Future Crime. *Journal of Human Resources, 47*(2), 509-549.

Currie, J., & Widom, C.S. (2010). Long-term consequences of child abuse and neglect on adult economic well-being. *Child Maltreatment, 15*(2), 111-120.

Currie, J., Stabile, M., Manivong, P., & Roos, L.L. (2010). Child health and young adult outcomes. *Journal of Human Resources, 45*(3), 517-548.

Dastan, I. (2011). Labor market effects of obesity, smoking, and alcohol use. *Dissertation Abstracts International, 72*(03), A.

Dee, T.S., & Evans, W.N. (2003). Teen drinking and educational attainment: Evidence from two-sample instrumental variables estimates. *Journal of Labor Economics, 21*(1), 178-209.

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.

Eckenrode, J., Laird, M., & Doris, J. (1993). School performance and disciplinary problems among abused and neglected children. *Developmental Psychology, 29*(1), 52-83.

Elkins, J.S., Whitmer, R.A., Sidney, S., Sorel, M., Yaffe, K., & Johnston, S.C. (2006). Midlife obesity and long-term risk of nursing home admission. *Obesity, 14*(8), 1472-8.

Ellickson, P.L., Tucker, J.S., & Klein, D.J. (2003). Ten-year prospective study of public health problems associated with early drinking. *Pediatrics, 111*(5), 949-55.

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.

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-Gők, V. (2003). *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., 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., & Horwood, L.J. (1997). Early onset cannabis use and psychosocial adjustment in young adults. *Addiction, 92*(3), 279-296.

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.

Fergusson, D.M., & Woodward, L.J. (2002). Mental health, educational, and social role outcomes of adolescents with depression. *Archives of General Psychiatry, 59*(3), 225-231.

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.

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., & Wolfe, B. (2009). Long-term consequences of childhood ADHD on criminal activities. *The Journal of Mental Health Policy and Economics, 12*(3), 119-138.

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.M., & Wolfe, B.L. (2009). 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.

Francesconi, M. (2008). Adult outcomes for children of teenage mothers. *The Scandinavian Journal of Economics, 110*(1), 93-117.

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.

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.

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.

Green, K.M., Doherty, E.E., Zebrak, K.A., & Ensminger, M.E. (2011). Association between adolescent drinking and adult violence: evidence from a longitudinal study of urban African Americans. *Journal of Studies on Alcohol and Drugs, 72*(5), 701-10.

Green, K.M., Doherty, E.E., Stuart, E.A., & Ensminger, M.E. (2010). Does heavy adolescent marijuana use lead to criminal involvement in adulthood? Evidence from a multiwave longitudinal study of urban African Americans. *Drug and Alcohol Dependence, 112,* 117-125.

Han, E., Norton, E.C., & Stearns, S.C. (2009). Weight and wages: fat versus lean paychecks. *Health Economics, 18*(5), 535-548.

Harris, Y., & Cooper, J.K. (2006). Depressive Symptoms in Older People Predict Nursing Home Admission. *Journal of the American Geriatrics Society, 54*(4), 593-597.

Hawkins, R., Jaccard, J., & Needle, E. (2013). Nonacademic factors associated with dropping out of high school: Adolescent problem behaviors. *Journal of the Society for Social Work and Research, 4*(2), 58-75.

Hill, K.G., White, H.R., Chung, I.J., Hawkins, J.D., & Catalano, R.F. (2000). Early adult outcomes of adolescent binge drinking: person- and variable-centered analyses of binge drinking trajectories. *Alcoholism, Clinical and Experimental Research, 24*(6), 892-901.

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.

Huang, S., Trapido, E., Fleming, L., Arheart, K., Crandall, L., French, M., Malcolm, S., ... Prado, G. (2011). *The long-term effects of childhood maltreatment experiences on subsequent illicit drug use and drug-related problems in young adulthood. Addictive Behaviors*, 36.

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.

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.

Kahn, M. E. (1998). Health and Labor Market Performance: The Case of Diabetes. *Journal of Labor Economics, 16*(4), 878-899.

Kane, J. B., Morgan, S. P., Harris, K. M., & Guilkey, D. K. (2013). The Educational Consequences of Teen Childbearing. Demography, 50, 6, 2129-2150.

Kazemian, L., Widom, C.S., Farrington, D.P. (2011) A prospective examination of the relationship between childhood neglect and juvenile delinquency in the Cambridge Study in Delinquent Development. *International Journal of Child, Youth and Family Studies, 1 & 2*, 65-82.

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.

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., 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.

Legleye, S., Obradovic, I., Janssen, E., Spilka, S., Le, N.O., & Beck, F. (2010). Influence of cannabis use trajectories, grade repetition and family background on the school-dropout rate at the age of 17 years in France. *The European Journal of Public Health, 20*(2), 157-163.

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.

Liu, S., Heaman, M., Kramer, M. S., Demissie, K., Wen, S.W., Marcoux, S., & Maternal Health Study Group of the Canadian Perinatal Surveillance System. (2002). Length of hospital stay, obstetric conditions at childbirth, and maternal readmission: a population-based cohort study. *American Journal of Obstetrics and Gynecology, 187*(3), 681-7.

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.

Luppa, Melanie, Luck, Tobias, Matschinger, Herbert, König, Hans-Helmut, & Riedel-Heller, Steffi G. (2010). Predictors of nursing home admission of individuals without a dementia diagnosis before admission—results from the Leipzig Longitudinal Study of the Aged (LEILA 75+). (BioMed Central Ltd.) BioMed Central Ltd.

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-Gők, 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.

Massetti, G. M., Lahey, B. B., Pelham, W. E., Loney, J., Ehrhardt, A., Lee, S. S., & Kipp, H. (2008). Academic achievement over 8 years among children who met modified criteria for Attention-deficit/Hyperactivity Disorder at 4-6 years of age. *Journal of Abnormal Child Psychology, 36*(3), 399-410.

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.

McCarren, M., Janes, G.R., Goldberg, J., Eisen, S.A., True, W.R., & Henderson, W.G. (1995). A twin study of the association of post-traumatic stress disorder and combat exposure with long-term socioeconomic status in Vietnam veterans. *Journal of Traumatic Stress, 8*(1), 111-124.

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.

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.

Minor, T. (2013). An investigation into the effect of type I and type II diabetes duration on employment and wages. *Economics and Human Biology, 11*(4), 534-544.

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. (1996). Employment, unemployment, and problem drinking. *Journal of Health Economics, 15*(4), 409-434.

Murray, J., Irving, B., Farrington, D. P., Colman, I., & Bloxsom, 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.

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.

Newcomb, M.D., & McGee, L. (1989). Adolescent alcohol use and other delinquent behaviors: One-year longitudinal analysis controlling for sensation seeking. *Criminal Justice and Behavior, 16*(3), 345-369.

Ng, Y.C., Jacobs, P., & Johnson, J. . (2001). Productivity losses associated with diabetes in the US. *Diabetes Care, 24*(2), 257-61.

Noll, J.G., Zeller, M.H., Trickett, P.K., & Putnam, F.W. (2007). Obesity risk for female victims of childhood sexual abuse: A prospective study. *Pediatrics, 120*(1), e61-e67.

Ou, S.-R., & Reynolds, A.J. (2010). Grade retention, postsecondary education, and public aid receipt. *Educational Evaluation and Policy Analysis, 32*(1), 118-139.

Peng, L., Meyerhoefer, C.D., Zuvekas, S., & National Bureau of Economic Research. (2013). *The effect of depression on labor market outcomes.* Cambridge, Mass: National Bureau of Economic Research.

Popovici, I., Homer, J.F., Fang, H., & French, M.T. (2012). Alcohol Use and Crime: Findings from a Longitudinal Sample of U.S. Adolescents and Young Adults. *Alcoholism: Clinical and Experimental Research,* 36(3), 532-543.

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.

Power, C., Pinto, P.S.M., & Li, L. (2015). Childhood Maltreatment and BMI Trajectories to Mid-Adult Life: Follow-Up to Age 50y in a British Birth Cohort. *Plos One, 10*, 3.

Renna, F. (2008). Teens' alcohol consumption and schooling. *Economics of Education Review, 27*(1), 69-78.

Resnick, S.G., & Rosenheck, R.A. (2008). Posttraumatic stress disorder and employment in veterans participating in Veterans Health Administration Compensated Work Therapy. *Journal of Rehabilitation Research & Development :A Publication of the Rehabilitation Research and Development Service, Department of Medicine and Surgery, Veterans Administration, 45*(3), 427-436.

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.
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.
Shenk, C.E., Putnam, F.W., Rausch, J.R., Peugh, J.L., & Noll, J.G. (2014). A longitudinal study of several potential mediators of the relationship between child maltreatment and posttraumatic stress disorder symptoms. *Development and Psychopathology, 26*(1), 81-91.
Shin, S.H., & Miller, D.P. (2012). A longitudinal examination of childhood maltreatment and adolescent obesity: Results from the National Longitudinal Study of Adolescent Health (AddHealth) Study. *Child Abuse & Neglect, 36*(2), 84-94.
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.
Songer, T.J., LaPorte, R.E., Dorman, J. S., Orchard, T.J., Becker, D.J., & Drash, A.L. (1989). Employment spectrum of IDDM. *Diabetes Care, 12*(9), 615-622.
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.
Stewart, W.F., Ricci, J.A., Chee, E., Hirsch, A.G., & Brandenburg, N.A. (2007). Lost productive time and costs due to diabetes and diabetic neuropathic pain in the US workforce. *Journal of Occupational and Environmental Medicine / American College of Occupational and Environmental Medicine, 49*(6), 672-679.
Stineman, M G., Xie, D., Streim, J.E., Pan, Q., Kurichi, J.E., Henry-Sanchez, J.T., Zhang, Z., ... Saliba, D. (2012). Home accessibility, living circumstances, stage of activity limitation, and nursing home use. *Archives of Physical Medicine and Rehabilitation,* 93(9), 1609-1616.
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? *Childrens 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., 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.
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.
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.
Tunceli, K., Bradley, C. J., Nerenz, D., Williams, L. K., Pladevall, M., & Elston, L. J. (2005). The impact of diabetes on employment and work productivity. *Diabetes Care, 28*(11), 2662-7.
Tunceli, K., Li, K., & Williams, L.K. (2006). Long-term effects of obesity on employment and work limitations among U.S. adults, 1986 to 1999. *Obesity, 14*(9), 1637-1646.
Valiyeva, E., Russell, L.B., Miller, J.E., & Safford, M.M. (2006). Lifestyle-related risk factors and risk of future nursing home admission. *Archives of Internal Medicine, 166*(9), 985-90.
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., & Williams, J. (2009). Why parents worry: Initiation into cannabis use by youth and their educational attainment. *Journal of Health Economics, 28*(1), 132-142.
Viner, R.M., & Taylor, B. (2007). Adult outcomes of binge drinking in adolescence: findings from a UK national birth cohort. *Journal of Epidemiology and Community Health, 61*(10), 902-7.
Webbink, D., Koning, P., Vujic, S., & Martin, N.G. (2012). Why are criminals less educated than non-criminals? Evidence from a Cohort of Young Australian Twins. *Journal of Law, Economics, and Organization*, 2.

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., Vujić, 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

Wells, J.E., Horwood, L J., & Fergusson, D.M. (2004). Drinking patterns in mid-adolescence and psychosocial outcomes in late adolescence and early adulthood. *Addiction, 99*(12), 1529-1541.

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.

WSIPP. (2013). *Analysis of crime and incarceration in Washington State*. Olympia: Washington State Institute for Public Policy.

WSIPP. (2014). *WSIPP analysis*. Olympia: Washington State Institute for Public Policy.

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.

Yan, J., & Brocksen, S. (2013). Adolescent risk perception, substance use, and educational attainment. *Journal of Risk Research, 16*(8), 1037-1055.

Zatzick, D.F., Marmar, C.R., Weiss, D.S., Browner, W.S., Metzler, T.J., Golding, J.M., Stewart, A., ... Wells, K.B. (1997). Posttraumatic stress disorder and functioning and quality of life outcomes in a nationally representative sample of male Vietnam veterans. *The American Journal of Psychiatry, 154*(12), 1690.

Zuvekas, S, Cooper, P.F., & Buchmueller, T.C. (2005). *Health behaviors and labor market status: The impact of substance abuse* (Working Paper No. 05013). Rockville, MD: Agency for Healthcare Research and Quality.

For further information, contact:
Stephanie Lee at 360.586.3951, stephanie.lee@wsipp.wa.gov

Washington State Institute for Public Policy

The Washington State Legislature created the Washington State Institute for Public Policy in 1983. A Board of Directors—representing the legislature, the governor, and public universities—governs WSIPP and guides the development of all activities. WSIPP's mission is to carry out practical research, at legislative direction, on issues of importance to Washington State.