Benefit-Cost Technical Documentation

Washington State Institute for Public Policy Benefit-Cost Model

December 2019

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 Michael Hirsch at michael.hirsch@wsipp.wa.gov or (360) 664-9081.

Washington State Institute for Public Policy
110 Fifth Avenue Southeast, Suite 214 • PO Box 40999 • Olympia, WA 98504-0999
(360) 664-9800 • 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.

Suggested citation: Washington State Institute for Public Policy. (December 2019). *Benefit-cost technical documentation.* Olympia, WA: Author.

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 State 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 State 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 programs and practices in the following public policy areas:

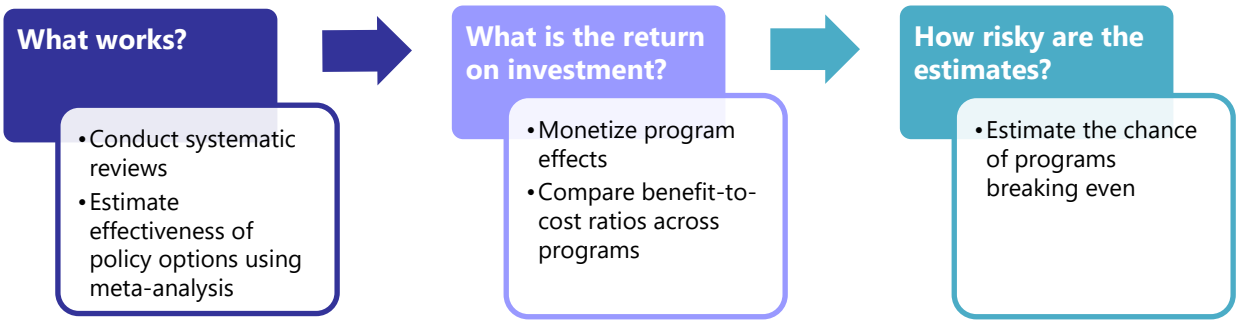
- ✓ Criminal and juvenile justice
- ✓ K–12 and early education
- ✓ Child welfare
- ✓ Substance abuse
- ✓ Mental health
- ✓ Public health
- ✓ Public assistance
- ✓ Employment and workforce development
- ✓ Health care
- ✓ General prevention
- ✓ Higher education

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.

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 (such as 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 three 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 U.S. (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 is the return on investment?** The second step involves applying economic calculations to put a monetary value on any changed outcome (from the first step). Once monetized, the estimated benefits are then compared to the costs of programs or policies to produce an economic bottom line for the investment. Chapters 4 and 5 describe 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 (from the second step) will at least break even. For example, if we conclude that on average, an investment in program XYZ has a ratio of $3 of benefits for each $1 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?

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 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 the 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.

$$(1.1.1)\quad \mathrm{NPV}_{\mathrm{Tage}} = \sum_{\mathrm{y}=\mathrm{Tage}}^{\mathrm{N}} \frac{\mathrm{Q_y} \times \mathrm{P_y} - \mathrm{C_y}}{(1+\mathrm{Dis})^{\mathrm{y}}}$$

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 1.1.1, 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 1.1.1. 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 1.1.1, a benefit-to-cost ratio, B/C, can be computed with:

$$(1.1.2) \quad \frac{B}{C} = \sum_{y=tage}^{N} \frac{Q_y \times P_y}{(1+Dis)^y} \Big/ \sum_{y=tage}^{N} \frac{C_y}{(1+Dis)^y}$$

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 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 exclude 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 can find on Functional Family Therapy. We do this through an "effect size", a statistical tool that allows for the combination of outcomes that have been measured in different ways.

"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 Appendices I and II of this document.

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 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. However, these two outcomes are likely to be, at least in part, measures of the same development in a person's human capital, with both leading 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: 1) the benefits and costs that accrue solely to program participants, 2) those received by taxpayers, 3) those received by others, and 4) those that are more indirect.

We created the third and fourth categories ("Others" and "Indirect," respectively) to report results that do not fit neatly in the first and second categories ("Participant" or "Taxpayer"). 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's 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. The evidence on effective public policy 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.[1]

With assistance from the Pew Charitable Trusts (Pew) and the MacArthur Foundation, WSIPP's benefit-cost model is being implemented in 20 other states and 11 county governments as part of the Pew-MacArthur Results First Initiative.[2] As part of our work with these organizations, the benefit-cost model has been reviewed four times in the past eight years by an independent team assembled by Pew. Most recently, the benefit-cost model was reviewed in 2017 by:

- ✓ D. Max Crowley: Assistant Professor of Human Development and Family Studies, Pennsylvania State University,
- ✓ Lynn Karoly: Senior Economist, Rand Corporation and Professor, Pardee RAND Graduate School
- ✓ David Weimer: Edwin E. Witte Professor of Political Economy, Robert M. La Follette School of Public Affairs, University of Wisconsin-Madison
- ✓ Frederick J. Zimmerman: Professor, Fielding School of Public Health, UCLA

The benefit-cost model was also reviewed in 2014 by Max Crowley, Lynn Karoly, David Weimer, and Paula Worthington (Senior Lecturer, Harris School of Public Policy, University of Chicago), 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 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 the 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.

[1] See: Drake, E. (2012). Reducing crime and criminal justice costs: Washington State's evolving research approach. *Justice Research and Policy, 14*(1), 97-116; 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 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.

[2] See http://www.pewtrusts.org/en/projects/pew-macarthur-results-first-initiative.

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 2.0.1.

$$(2.0.1)\quad NPV_{tage} = \sum_{y=tage}^{N} \frac{Q_y \times P_y - C_y}{(1+Dis)^y}$$

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 2.0.1, Q_y, is the estimated quantity of outcome "units" in year y produced by the program or policy. The Q_y term in Equation 2.0.1 is, in turn, a function of two factors in the WSIPP model: an "effect size" (*ES*) and a "*Base*" variable as given by Equation 2.0.2.

$$(2.0.2)\quad Q_y = f(ES, Base)$$

The "effect size" is a statistical method to compare the relative magnitude of effects. The *Base* variable is the amount of the outcome in the targeted population without the intervention. Examples include the proportion of a criminal justice population that is expected to commit another crime or the proportion of teens expected to give birth in the absence of intervention.

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). Using the effect size measure allows for this combination of different measures.

For continuously measured outcomes, as given by Equation 2.0.3 and described later in this chapter and in Chapter 3, Q_y is calculated with a Cohen's d (standardized mean difference) effect size[3] and a *Base* variable, which is measured as a standard deviation of the outcome measurement.

$$(2.0.3)\quad Q_y = Base_y \times ES$$

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 2.0.4.[4]

$$(2.0.4)\quad Q_y = \frac{\left(e^{ES \times 1.65} \times Base_y\right)}{\left(1 - Base_y + Base_y \times e^{ES \times 1.65}\right)} - Base_y$$

[3] Lipsey, M.W., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

[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.

Exceptions. Two of the exceptions to this equation for estimating Q_y for continuously measured outcomes are 1) when 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 as well as health care costs and frequency of visits measured by evaluations of certain health care programs), and 2) when 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 2.0.5 below.

$$(2.0.5)\ Q_y = ES$$

Another exception to this equation occurs when an outcome is measured as an incidence rate ratio (currently used to value changes to the fall rate among older adults). This ratio is applied using Equation 2.0.6 below.

$$(2.0.6)\ Q_y = Base_y \times ES - Base_y$$

This chapter describes the process we use to estimate the effect size term, ES, in Equations 2.0.3 to 2.0.6. Chapter 3 discusses how Q_y is then estimated from the effect sizes and dichotomous or continuous base variables. In Chapter 4 we describe the various methods we use to estimate the price term, P_y, in equation 2.0.1. In Chapter 6 we describe our procedures for computing program costs, C_y, in Equation 2.0.1.

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

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.

However, the outcome evaluations of FFT did not measure whether the program affects high school graduation rates—another outcome of keen interest to the Washington State 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 Appendix I.

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, 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] The following are the key criteria we implement when conducting a meta-analysis.

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. Non-peer reviewed studies also represent a significant portion of the available evidence in many policy areas. 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 or pre-treatment characteristics such as age, gender, test scores, or level of functioning.

[5] In general, we follow the meta-analytic methods described in Lipsey & Wilson (2001).
[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. Many other studies were reviewed but did not meet the criteria set for this analysis.

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 in Section 2.3. If the necessary information is not provided, and we are unable to obtain it directly from the study's author(s), the study is not included in our review.

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 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, rather than exact values. Examples include: $p < 0.01$, $p < 0.05$, or non-significant at the $p \geq 0.05$ 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. Authors report outcomes in different ways depending on the research design and the nature of the outcomes. For example, an author may report the number of participants who report remaining sober after participating in a substance treatment program. Alternatively, the authors may have information on the results of drug screens. Our goal is to simplify the diversity of outcome statistics reported across multiple studies into a single measure: the effect size. We also calculate the variance around the effect size for each outcome, for two reasons. We use the variance of outcomes from individual studies to calculate weighted average effect sizes as discussed in Section 2.3e. Additionally, we use the variation around the weighted average effect size in our Monte Carlo simulation as described in Chapter 7.

Analysts use several methods to calculate effect sizes, as described in Lipsey & Wilson.[8] The most common effect size statistics (and the measures we use in our meta-analyses) are the standardized mean difference effect size for continuous outcomes (Section 2.3a) and the Cox transformation of a dichotomous variable to the standardized mean difference effect size (Section 2.3b). In special circumstances, we will also perform a meta-analysis on elasticities, semi-elasticities, and incidence rate ratios (Section 2.3c).

[7] Lipsey & Wilson (2001).
[8] Ibid.

2.3a Continuously Measured Outcomes

The mean difference effect size is designed to accommodate continuous outcome data, such as student test scores, where the differences are between the means of the outcome.[9] The standardized mean difference effect size is computed with the following equation:

$$(2.3.1)\quad ES = \frac{M_t - M_c}{\sqrt{\frac{(N_t - 1)SD_t^2 + (N_c - 1)SD_c^2}{N_t + N_c - 2}}}$$

In this formula, *ES* is the estimated effect size for a particular program; M_t is the mean value of an outcome for the treatment or experimental group; M_c is the mean value of an outcome for the control group; SD_t is the standard deviation of the treatment group; and SD_c is the standard deviation of the control group; N_t is the number of subjects in the treatment group; and N_c is the number of subjects in the control group. In instances where there is insufficient information to determine the division of subjects between treatment and control, we assume an equal division of the total N.

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:[10]

$$(2.3.2)\quad ES = t\sqrt{\frac{N_t + N_c}{N_t N_c}}$$

We compute the variance of the mean difference effect size statistic in Equation 2.3.1 with the following equation:[11]

$$(2.3.3)\quad ESVar = \frac{N_t + N_c}{N_t N_c} + \frac{ES^2}{2(N_t + N_c)}$$

2.3b 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. show 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:

$$(2.3.4)\quad ES_{Cox} = \frac{\ln\left[\frac{P_t(1 - P_c)}{P_c(1 - P_t)}\right]}{1.65}$$

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 the following variance:

$$(2.3.5)\quad ESVar_{Cox} = 0.367\left[\frac{1}{O_{1t}} + \frac{1}{O_{2t}} + \frac{1}{O_{1c}} + \frac{1}{O_{2c}}\right]$$

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.

[9] Lipsey & Wilson (2001), table B10, equation 1, p. 198.
[10] Ibid, table B10, equation 2, p. 198.
[11] Ibid, table 3.2, p. 72.
[12] Sánchez-Meca et al. (2003).

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] and then multiply the result by 1.35 to determine ES_{Cox} as given by the following equation:

$$(2.3.6)\quad |ES_{cox}| = 1.35 * 2\sqrt{\frac{X^2}{N_t + Nc - X^2}}$$

Similarly, we determine that in these cases using Equation 2.3.2 to calculate the variance underestimates $ESVar_{Cox}$ and hence overestimates 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 $ESVar_{Cox}$.

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 2.3.4, i.e. by taking the log of the odds ratio, divided by 1.65.

The variance is calculated using the following equation:

$$(2.3.7)\quad ESVar_{Cox} = 0.367 * \left(\frac{\frac{\ln(upper\ CI) - \ln(lower\ CI)}{2}}{1.96}\right)^2$$

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.

2.3c Other Effect Size Methods

In addition to calculations for regular measurements of continuously measured variables using the standardized mean difference, we have special calculation rules for elasticities and semi-elasticities and for incidence rate ratios. These special calculations cannot be combined with our typical standardized mean difference approach.

Effect Sizes Measured as Elasticities or Semi-elasticities. Some areas of research we review tend to take an econometric approach; that is, studies use regression techniques 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. Another common metric is 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 bodies of research 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 a high deductible health care plan on health care costs often use semi-elasticities estimated as a log-linear model. We would then estimate a semi-elasticity, or percent change, in health care costs due to participation in a high-deductible plan by exponentiating the β from the regression and subtracting one to calculate the percent change. 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.

For effect sizes measured as elasticities, the SEe is equivalent to the standard error of the elasticity. When a study reports the standard error on the elasticity, we use that value as SE. The standard error of the elasticity is most commonly reported when the study estimates the elasticity from a log-log 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 from the linear model using the following equations.

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

[13] Lipsey & Wilson (2001), table B10, equation 23, p. 200.

$$(2.3.8)\ Var(Elas) = \frac{X^2}{Y^2}Var(\beta_1) + \beta_1^2 * \frac{X^2}{Y^4} * Var(Y)$$

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

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

$$(2.3.9)\ Var(SemiElas\ or\ \%\ Change) = \left(\frac{Y_t}{Y_c}\right)^2 * \left(\frac{Var(Y_c)}{Y_c^2} + \frac{Var(Y_t)}{Y_t^2}\right)$$

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 or in the case of a semi-elasticity from a log-linear model, **,** we assume that the elasticity or semi-elasticity has the same statistical significance as the regression coefficient from which we derive the elasticity or semi-elasticity. Under this assumption, 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:

$$(2.3.10)\quad Var(SemiElas\ or\ \%\ Change) \quad = \left(\frac{ES_\epsilon}{t_\beta}\right)^2.$$

Effect Sizes Measured as Incidence Rate Ratios. Occasionally we review research that reports count data as rates. Rates reflect a count of events for each individual over the observation period, often expressed in person-years. Analyzing count data as rates assumes a constant underlying risk of the event.

The preferred effect size for outcomes reported as rates is an incidence rate ratio (IRR), rather than the other effect size metrics (Cohen's *d* or D-cox effect sizes). The IRR is the ratio of the number of events per person-year among individuals in the intervention group to the number of events per person-year among individuals in the comparison group. IRRs have particular properties to consider when conducting a meta-analysis. For example, a "null" incidence rate ratio is one (not zero).

We use the methodology described by the Cochrane Collaboration aggregate incidence rate ratios.[14] We calculate the natural logarithm of the IRR as the effect size for a given study, and combine the natural logarithms of the IRR in our meta-analysis. We transform the results back to the linear scale and report results as incidence rate ratios on the linear scale for interpretability.

For incidence rate ratio effect sizes, we use an approximate standard error of the natural logarithm of the IRR to calculate the inverse variance weight. The equation for the approximate standard error is available from Cochrane,[15] and relies on the number of events in the treatment and comparison groups:

$$(2.3.11)\ Var(ln(rate\ ratio)) = \frac{1}{events_T} + \frac{1}{events_C}$$

2.3d Modifying Effect Sizes to Account for Small-Sample Sizes and Multi-Level Data Structures

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,"

[14] Deeks, J.J., Higgins, J.P.T., & Altman, D.G. (2011). Chapter 9: Analysing data and undertaking meta-analyses. In J.P.T. Higgins & S. Green (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011.
[15] Deeks et al. (2011).

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: [16]

$$(2.3.12)\quad ES'_m = \left[1 - \frac{3}{4N-9}\right] * \ ES_m$$

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 underestimate the variance in outcomes and, thus, may overestimate effect sizes. In studies that do not account for clustering, effect sizes and their variance require additional adjustments.[17]
We account for clustering differently for continuous and dichotomous outcomes. We do not currently make clustering adjustments on topics that include outcomes reported as elasticities, semi-elasticities, or incidence rate ratios.

Adjustments for clustering for continuous outcomes. We adjust based on whether the information is reported at the individual or the cluster level.[18]

First, for continuous outcomes in studies reported at the individual-level that ignore the variance due to clustering, we make adjustments to the uncorrected effect size and its variance, using the following equation:

$$(2.3.13)\quad ES_T = \ ES_u * \sqrt{1 - \frac{2(n-1)\rho}{N-2}}$$

$$(2.3.14)\quad Var(ES_T) = \left(\frac{N_t + N_c}{N_t N_c}\right)[1 + (n-1)\rho] + \ {ES_T}^2 \left(\frac{(N-2)(1-\rho)^2 + n(N-2n)\rho^2 + 2(N-2n)\rho(1-\rho)}{2(N-2)[(N-2) - 2(n-1)\rho]}\right)$$

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 data from the 2006 Washington Assessment of Student Learning (WASL) 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 continuous outcomes in 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:

$$(2.3.15)\quad ES_T = ES_m * \sqrt{\frac{1 + (n-1)\rho}{n\rho}} * \sqrt{\rho}$$

$$(2.3.16)\quad Var(ES_T) = \left\{\left(\frac{N_t - N_c}{N_t N_c}\right) * \left(\frac{1 + (n-1)\rho}{n\rho}\right) + \frac{[1 + (n-1)\rho]^2 * \ {ES_T}^2}{2n\rho(K-2)}\right\} * \rho$$

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 for outcomes that use continuous measures.

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

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

$$(2.3.17) \quad ES_{int} = ES * \sqrt{1 - \frac{m_t(n_t - 2)\rho}{N - 2}}$$

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.

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]
d

$$(2.3.18) \quad h = \frac{[(N-2)(\rho - 1) + (n_t m_t - n_t)\rho]^2}{(N-2)(1-\rho)^2 + (n_t m_t - n_t)n_t\rho^2 + 2(n_t m_t - n_t)\rho(1-\rho)}$$

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

$$(2.3.19) \quad J(h) = 1 - \frac{3}{4h - 1}$$

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

$$(2.3.20) \quad ES_T = ES_{int} * J(h)$$

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

$$(2.3.21) \quad ESVar = \frac{1 + (n-1)\rho}{n_t m_t} + \frac{1 - \rho}{m_c} + \frac{([N-2)(1-\rho)^2 + (n_t m_t - n_t)n\rho^2 + 2(n_t m_t - n)\rho(1-\rho] * ES_t}{2[(N-2)(1-\rho) + (n_t m_t - n_t)\rho]^2}$$

Adjustments for clustering for dichotomous outcome variances. We do not make a clustering adjustment to effect sizes in dichotomous outcomes. This is because the Cox transformation assumes the entire normal distribution at the student level.[24] However, when outcomes are dichotomous, we use the "design effect" to calculate the "effective sample size."[25] The effective sample size is used to calculate a corrected variance. The design effect is given by the following equation:

$$(2.3.22) \quad D = 1 + (n-1)\rho$$

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:

$$(2.3.23) \quad n_{t(eff)} = \frac{n_t}{D}$$

We recalculate the variance on these dichotomously measured effect sizes with the methods described in Section 2.3b, substituting $n_{t(eff)}$ for n_t and $n_{c(eff)}$ for n_c.

[19] Larry Hedges (personal communication, June 11, 2012).
[20] Ibid.
[21] Ibid.
[22] Ibid.
[23] Ibid.
[24] Mark Lipsey (personal communication, November 11, 2007).
[25] 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.

2.3e Computing Weighted Average Effect Size

Computing Weighted Average Effect Size and Standard Error. Once effect sizes are calculated for each program effect, and any necessary adjustments for clustering are made, the individual measures are used to produce a weighted average effect size for each outcome within the program. Each effect size is weighted by the inverse of the variance of the effect size, *ESVar*, as described in the preceding sections.

$$(2.3.24) \quad w_i = \frac{1}{ESVar_i}$$

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

$$(2.3.25) \quad \overline{ES} = \frac{\sum(w_i ES_{T_i})}{\sum w_i}$$

The standard error of this estimate is calculated with the following equation:[27]

$$(2.3.26) \quad SE_{\overline{ES}} = \sqrt{\frac{1}{\sum w_i}}$$

After computing the fixed effects weighted average effect size and standard error, we compute the random effects if necessary.

Computing Homogeneity Tests, Random Effects Weighted Average Effect Sizes, and Standard Error. 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.[28]

First, we test for homogeneity. 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]

$$(2.3.27) \; Q = \left(\sum w_i \, ES_i^2\right) - \frac{\left(\sum w_i \, ES_i^2\right)}{\sum w_i}$$

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

Next, we check whether there is adequate variation to use the Random Effects model. We proceed if the following is true:

$$(2.3.28) \; 0 < Q_i - (k-1)$$

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. If not, we calculate the random effects variance component, *v* using the following equation:[30]

$$(2.3.29) \; v = \frac{Q_i - (k-1)}{\sum w_i - \left(\sum wsq_i / \sum w_i\right)}$$

where wsq_i is the square of the weight of ES_i (Equation 2.3.15).

This random variance factor is added to the variance of each effect size and all inverse variance weights are recomputed as follows:

$$(2.3.30) \quad REw_i = \frac{1}{ESVar_i + v}$$

The Effect Size is recalculated using the REw_i:

$$(2.3.31) \quad \overline{ES} = \frac{\sum(REw_i ES_{T_i})}{\sum REw_i}$$

The variance is recalculated as:

$$(2.3.32) \quad SE_{\overline{ES}} = \sqrt{\frac{1}{\sum REw_i}}$$

---

[26] Lipsey & Wilson (2001), p. 114.
[27] Ibid.
[28] 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.
[29] Ibid, p. 116.
[30] Ibid. p. 134.

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. We call these effects "unadjusted effect sizes." In our reports and on our website, we also list an "adjusted effect size" for each outcome. These adjusted effect sizes are modifications of the unadjusted results. They may be smaller, larger, or equal to the unadjusted effect sizes we report. Importantly, 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 works to improve outcomes in Washington. If a program has been rigorously tried and tested somewhere else, we want to be able to infer whether it is likely to work in Washington. We believe there is reason to be concerned that the results of individual program evaluations (the ones we enter into our meta-analyses) may be different if the program were to be implemented in Washington. This is because 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. We may make adjustments to account for any of the following characteristics:

1) The methodological quality of each study we include in a meta-analyses;
2) Whether the researcher(s) who conducted a study is (are) invested in the program's design and results;
3) The relevance or quality of the outcome measured used in a study;
4) Whether the research was conducted in a laboratory or other unusual "non-real world" setting; and
5) Situations in which an evaluation of a program was conducted against a wait-list or no treatment comparison group, as opposed to a treatment-as-usual comparison group.

We do not currently make adjustments to effect sizes that are computed as elasticities, semi-elasticities, or incidence rate ratios as covered in Section 2.3c.

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, producing results that are 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 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 differences in the quality of research designs may, on average, affect a program's true effect on outcomes. 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.[31]
- **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 public service delivery system. There is some evidence that programs closely controlled by researchers or program developers have consistently better results than those that operate in "real-world" administrative structures.[32] Therefore, because we are concerned that effects observed in developer-controlled evaluations may often overstate the effects we might expect in a 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 the corresponding effect size(s) to reflect this distinction. 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 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 WSIPP's assignments from the Washington State Legislature is to identify programs that can make cost-beneficial improvements to Washington's 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. 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 utilize treatment as usual for the comparison group, other studies compare a treatment group to a wait-list or no-treatment comparison 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 or wait-list 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. We list our current adjustments for weak outcome measures in Section 2.4f.

---

[31] 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 and 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.

[32] 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.

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 apply 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.

To estimate these adjustment factors, we undertake a series of meta-regression analyses, one for each broad research area. In some cases, where the research literature is particularly large, we may perform 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. In topic areas where there is a clear primary outcome (for example, depression outcomes in interventions for child depression) we include only the effect sizes from primary outcomes in our meta-regression. In these cases, we do not cluster standard errors by each study in the analysis, because each study only contributes one effect size to the analysis.

Our independent variables typically include the previously discussed five factors. Adjustment factors (in the form of multipliers) may be assigned to the results of individual effect sizes based on our findings of persistent statistical significance ($p<0.10$) for coefficients across a number of specifications.

After considered technical review, WSIPP is moving forward with a new method for calculating and applying multiplicative adjustment factors. This section describes both our historical approach and our new approach. The new framework is being implemented across research areas as we update the literature and meta-analyses. To date, we have used the new approach for juvenile justice topics. All of the other research areas still rely on our historical approach. We will update these meta-regression analyses and multiplicative adjustment factors as time and resources allow.

Historical Approach—Separate Multiplicative Adjustment Factors: We have historically calculated adjustment factors within a research area using Equation 2.4.1, where β_f is the regression coefficient for factor f (researcher = developer, weak outcome measure, etc.), d_{if} is an indicator variable indicating the presence of each adjustment factor for each study, and α is the intercept. These coefficients typically came from a preferred specification estimating a linear regression on the effect size including random effects. The adjusted effect size is calculated using Equation 2.4.2, which incorporates each separate adjustment factor multiplicatively.

$$(2.4.1)\ Adjustment\ factor_{if} = \frac{\alpha}{\alpha+(\beta_f \times d_{fi})}$$

$$(2.4.2)\ ESadj_i = ES_i \times \prod_f (Adjustment\ factor_{if})$$

Using this approach, adjustments are made by multiplying the unadjusted effect size for each study by each of the relevant adjustment factors. Exhibit 2.4.1 lists the current multiplicative adjustment factors for research areas that rely on this approach. The resulting meta-analytic findings for the adjusted effect sizes are then used in the benefit-cost analysis, as explained in Section 2.6.

Exhibit 2.4.1

Current WSIPP Adjustments—
Separate Multiplicative Adjustment Factors Applied to Unadjusted Effect Sizes

Topic area	Multiplicative Adjustment Factor				
	Research design	Researcher = developer	Weak outcome measure	"Not real world"	Wait-list design
Adult criminal justice	Level 1 = 0.395 Level 4 = 0.365 All others = 1	1	1	0.50	n/a
Substance abuse prevention	1	0.33	1	1	1
Substance abuse treatment	1	1	1	1	1
Early childhood education	1	1	1	1	n/a
Child welfare	1	0.36	1	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
Child depression	1	1	1	1	0.31
Child anxiety	1	1	1	1	0.59
Child posttraumatic stress	1	1	1	1	0.50
Child disruptive behavior	1	0.52	0.54	1	0.44
Child ADHD	1	0.51	1	1	0.40
General prevention/public health	Level 1 =0.31 All others = 1	0.38	1	1	1
Asthma self-management education	1	0.36	0.5	1	1
Workforce development	1	1	1	1	1
Workforce development: training with work experience	Level 1 = 0.62 Level 2 = 0.93 All others = 1	1	1	1	1
Health care	1	1	1	1	1
K-12 education	1	0.43	0.23	0.22	n/a
Higher education	Level 1 = 0.53 Level 2 = 0.53	1	1	1	1

Note:
In cases in which the adjustment factor does not have a statistically significant ($p < 0.10$) coefficient on the adjustment factor included in the meta-regression, we do not estimate an adjustment. The multiplier is 1, representing no adjustment.

New Approach—Single Combined Adjustment Factor: After considered technical review, WSIPP is moving forward with a new method for calculating and applying multiplicative adjustment factors. We use the same model parameters for the meta-regression described above. With the new approach, we calculate a single combined multiplier (combined adjustment factor) for each study as shown in Equation 2.4.3 and apply it using Equation 2.4.4.

$$(2.4.3)\ Combined\ adjustment\ factor_i = \frac{\alpha}{\alpha + \sum_f (\beta_f \times d_{fi})}$$

$$(2.4.4)\ ESadj_i = ES_i \times Combined\ adjustment\ factor_i$$

This method still accounts for the fact that a single study may have multiple characteristics that require an adjustment. Instead of multiplying the effect size by each separate relevant multiplicative adjustment factor, this approach creates a single multiplier (combined adjustment factor) that accounts for each of relevant adjustment factors. Exhibit 2.4.2 lists the current coefficients for research areas that rely on this approach. The resulting meta-analytic results for the adjusted effect sizes are then used in the benefit-cost analysis, as explained in Section 2.6.

Exhibit 2.4.2

Current WSIPP Adjustments—
Coefficients Used To Calculate Combined Multiplicative Factors Applied to Unadjusted Effect Sizes

Topic area	Coefficients Used to Estimate Combined Multiplicative Adjustment Factors					
	Constant	Research design	Researcher = developer	Weak outcome measure	"Not real world"	Wait-list design
Juvenile justice	1	0	0	0	0	0

Note:
In cases in which the adjustment factor does not have a statistically significant ($p < 0.10$) coefficient on the adjustment factor included in the meta-regression, we do not estimate an adjustment. The coefficient is 0, representing no adjustment. In cases in which none of the adjustment factors were statistically significant, we use a constant of 1 and a coefficient of 0, which results in no adjustment.

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 2.3.2 to calculate the adjusted variance (Var_{ad}) substituting the adjusted effect size (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:

$$(2.4.5)\ OR_{adj} = e^{(1.65\ ES_{adj})}$$

Next, we calculate the corresponding treatment percentage, assuming the comparison rate does not change.
Finally, we calculate the variance per Equation 2.3.5 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 2.3.2 and ES_{adj}. Then, based on our analysis, we multiply the Var_{adj} by 1.65 to provide a good approximation of $Var_{adj}Cox$.

2.5 WSIPP Adjustments to Effect Sizes from Longitudinal Linkage Studies

As with the results from program evaluations (discussed in Section 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 to too few studies on which 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 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 Sections 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 2.3.30 through 2.3.32, substituting the WSIPP-adjusted effect sizes and adjusted inverse variance weights 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 2.0.1, 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 substance abuse treatment 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 the follow-up period and effect size. Depending on the research area and available information, we may either use only the longest follow-up from each study or use all follow-up periods from a given study.[33] 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 the 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[34]
- ✓ 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.[35]

Finally, in some cases, we are unable to estimate program effects beyond the first effect size using either meta-analysis or regression analysis. This may occur 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 may 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 may assume that effects decay to zero for all time periods following those measured in the studies.

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

Exhibit 2.7.1

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	0	0.187	Age at Time 1 + 3
Contingency management (higher-cost)	0	0.125	Age at Time 1 + 1
Contingency management (lower-cost)	0	0.075	Age at Time 1 + 1
Medication-assisted therapies	0	0	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	0.141	Age at Time 1 + 1
Child depression	0	0.310	Age at Time 1 + 2
Child anxiety	ES1* 0.396	$(ES1^2 * 0.276^2 + 0.396^2 * SE1^2 + SE1^2 * 0.276^2)^{0.5}$	Age at Time 1 + 1
Child internalizing	ES1	SE1	Age at Time 1 + 2
Child externalizing, child disruptive behavior	ES1 * 0.550	$(ES1^2 * 0.550^2 + 0.238^2 * SE1^2 + SE1^2 * 0.550^2)^{0.5}$	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	0	0.054	Age at Time 1 + 7
Short-term diabetes interventions	ES1 * 0.31	0.101	Age at Time 1 + 7
Obesity prevention for children	0	0.070	Age at Time 1 + 2
Obesity prevention, adults, high-intensity	0	0.012	Age at Time 1 + 5
Obesity prevention, adults, low-intensity	0	0.012	Age at Time 1 + 2
Obesity			
Obesity prevention for children	0	0.101	Age at Time 1 + 2
Obesity prevention, adults, high-intensity	0	0.086	Age at Time 1 + 5
Obesity prevention, adults, low-intensity	0	0.086	Age at Time 1 + 2
Emergency room visits for asthmatic children or general population	0	0.086	Age at Time 1 + 2
Hospitalizations (readmissions) Patient-centered medical homes Outcomes for seriously mentally ill individuals, those easily lost to follow up Birth outcomes Falls	0	0	Age at Time 1 + 1
Labor market earnings (measured directly)			
Case management programs	0	0.014	Age at Time 1 + 1
Job search and placement	0	0.017	Age at Time 1 + 2
Training, no work experience	0	0.032	Age at Time 1 + 1
Training with work experience	0	0.018	Age at Time 1 + 1
Work experience	0	0.001	Age at Time 1 + 2

Notes:
Figures have been rounded to three decimal places.
ES1 = effect size at time 1. This is the effect size reported in the study at a follow-up time (usually 1-2 years after the intervention).
Time 2 = a time in the future after time 1. These vary by outcome.
SE = standard error.

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 "outcome effect sizes" rather than simply "outcome effects." This seemingly arcane distinction is important for our approach to benefit-cost modeling. Some important concepts are defined below.

- ✓ **"Outcome 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 four percentage points—from, say, 75% without the program to 79% with the program. This is an outcome 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.
- ✓ **"Outcome 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 an outcome effect into an effect size metric with the procedures described in Chapter 2 to allow us to combine the outcome effect with other outcome effects that might be measured differently. 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 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.
- ✓ **"Unit Change."** For purposes of clarity in this presentation, we call a program's effect on an outcome a "unit change" 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 2.3.4, of +0.137 for the four percentage point program effect (increase in high school graduation) in the hypothetical program evaluation. At this point, we have the following evidence from a single study:

- ✓ Percentage change for graduation rate from a single study: +4%
- ✓ Effect size for graduation rate from a single study: +0.137

We would then make similar effect size calculations for all of the tutoring studies in our meta-analysis and might conclude, for example, that tutoring programs, on average, can be expected to have a D-cox effect size of +0.15 on high school graduation. At this point, we have the following evidence from a meta-analysis:

- ✓ Effect size for graduation rate from all studies in the meta-analysis: +0.15

From this effect size finding, 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, 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, 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.

Finally, in some instances, we elect not to monetize certain outcomes in a specific meta-analysis. There are a few common scenarios in which we might elect not to monetize particular outcomes. These include:

- ✓ The outcome is measured in a single study with a small number of individuals or a limited or non-representative sample;
- ✓ WSIPP does not have an appropriate population in the model to monetize a particular outcome (for example, if the outcome is only measured in a high-risk population, but WSIPP's model only has the capability to model a "general" population for that outcome); or
- ✓ The meta-analysis has several outcomes measured in multiple studies and some outcomes that are measured in only one study, which has a limited or non-representative sample.

In these cases, WSIPP may only report the program effect sizes from the meta-analysis. These instances are noted in the meta-analysis tables on our website.

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, among others. 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 (*Base)* about the population being served by a program. This is given in the following equation:

$$(3.1.1)\quad Q_y = f(ES, Base)$$

In the WSIPP benefit-cost model, Equation 3.1.1 is operationalized with several user-supplied parameters. For each topic for which a benefit-cost analysis is to be calculated, these eight 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*, used in Monte Carlo draws
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*, used in Monte Carlo draws
Base	estimated outcome for the non-treatment group (i.e., 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. A single measured outcome may have more than one *Base*. For example, a program may be targeted towards those receiving treatment for alcohol use disorder. We expect these people to have a higher incidence (base rate) of alcohol use disorder than a program directed to the population at large. In these cases, the user is able to select a target population from a list of choices, thus populating the *Base* with the appropriate estimate.

The user first enters *Tage,* the age when the first program effect for a given outcome was measured, and *Mage1*, the first measurement age. If the user has conducted a meta-analysis, *Mage1* 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 enters the age of the person treated when a second program effect was measured or projected, *Mage2*. *Mage2* will always be greater than *Mage1*; it is designed as a way to project the longer-term effectiveness of a program. Program effects could decay, grow, or stay the same as time passes following treatment. The second follow-up period allows us to model the trajectory of these longer-term 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 Section 2.7, WSIPP increasingly conducts meta-regressions to inform our projection of longer-term program effect sizes.

For example, in a previous examination of the literature for the juvenile justice program called Functional Family Therapy (FFT), the assumed treatment age for the average juvenile in this program was 15. Next, the user input six of the eight parameters for the crime outcome measured for FFT. The first effect size was -0.247 and had a standard error of 0.120. For this program, our review of the FFT evaluations indicated that the average follow-up period was about two years; thus, we entered age 17 as *Mage1*. The second effect size, -0.247, was entered for age 27 with a standard error of 0.120. In the case of juvenile justice programs, the longer-term outcome was the same as that entered at the first follow-up period because our meta-regressions have indicated that effects of programs on crime 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 score effects decay over time. The WSIPP model accommodates the modeling of these time-dimensioned outcomes with this two-point process.

For each outcome represented in a meta-analysis, the user selects an appropriate population for that 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 graduation 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 (i.e., input by the user) and entered into the model, 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 (base rate).

3.2a Continuously Measured Outcomes

When outcomes are continuous, as given by Equations 3.2.1 and 3.2.2, the change in units at the first and second measurement ages, *Mage1* and *Mage2*, is calculated with a Cohen's d effect size and a $Base$ variable, which is measured as a standard deviation of the outcome measurement.

$$(3.2.1)\quad Q_{Mage1} = Base_1 \times ES_1$$

$$(3.2.2)\quad Q_{Mage2} = Base_2 \times ES_2$$

1) We distribute the unit change calculated at $Mage1$ (Equation 3.2.1) to the ages between $Tage$ and $Mage1$.
2) We distribute the unit change calculated at $Mage2$ (Equation 3.2.2) 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 3.2.1 and 3.2.2 are implemented using random draws from a normal probability density distribution centered on the unit change (Q_{mage1} and Q_{mage2}). The standard error of the normal distribution is calculated as the unit change multiplied by the coefficient of variation at that point (Equations 3.2.3 and 3.2.4). A common randomly drawn seed is used to compute both Q_{mage1} and Q_{mage2} for each Monte Carlo run.

$$(3.2.3)\quad Q_{seMage1} = Q_{Mage1} \times ES_{se}1/ES1$$

$$(3.2.4)\quad Q_{seMage2} = Q_{Mage2} \times ES_{se}2/ES2$$

Applications of Continuously Measured Outcomes for Non-Cohen's d Effect Size Measurements. Elasticities and semi-elasticities, as described in Section 2.3c, are calculated similarly to the Cohen's *d* effect size.

Incidence rate ratios, as described in Section 2.3c, have a slightly different calculation method.

$$(3.2.5)\quad Q_{Mage1} = Base_1 - (IRR_1 \times Base_1)$$

$$(3.2.6)\quad Q_{Mage2} = Base_2 - (IRR_2 \times Base_2)$$

1) We distribute the unit change calculated at *Mage1* (Equation 3.2.5) to the ages between *Tage* and *Mage1*.
2) We distribute the unit change calculated at *Mage2* (Equation 3.2.6) to the age at *Mage2*.
3) For ages ranging from *Mage1* to *Mage2*, we linearly interpolate the unit change between *Mage1* and *Mage2*.
4) For ages greater than *Mage2*, we set the unit change to 0.

In Monte Carlo simulations, Equations 3.2.5 and 3.2.6 are implemented using random draws from a normal probability density distribution centered on the unit change (Q_{mage1} and Q_{mage2}). The standard error of the normal distribution is calculated as the unit change multiplied by the coefficient of variation at that point (Equations 3.2.7 and 3.2.8). The coefficient of variation calculated at *Mage1* ($ES_{se}1/(1-ES1)$) is applied to all ages from *Tage* to the age prior to *Mage2*. A common randomly drawn seed is used to compute both Q_{mage1} and Q_{mage2} for each Monte Carlo run.

$$(3.2.7)\quad Q_{seMage1} = Q_{Mage1} \times ES_{se}1/(1-ES1)$$

$$(3.2.8)\quad Q_{seMage2} = Q_{Mage2} \times ES_{se}2/(1-ES2)$$

3.2b Dichotomously Measured Outcomes

As given by Equations 3.2.9 and 3.2.10 below, 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 3.2.1 provides a numeric example to illustrate these procedures for dichotomous outcomes, which are slightly more complex than the procedure for continuous outcomes.

$$(3.2.9)\quad Q_{Mage1} = \left(\frac{(e^{ES_1 \times 1.65} \times Base_1)}{(1 - Base_1 + Base_1 \times e^{ES_1 \times 1.65})} - Base_1\right)$$

$$(3.2.10)\quad Q_{Mage2} = \left(\frac{(e^{ES_2 \times 1.65} \times Base_2)}{(1 - Base_2 + Base_2 \times e^{ES_2 \times 1.65})} - Base_2\right)$$

$$(3.2.11)\quad Q_{seMage1} = Q_{Mage1} \times ES_{se}1/ES1$$

$$(3.2.12)\quad Q_{seMage2} = Q_{Mage2} \times ES_{se}2/ES2$$

- ✓ Equations 3.2.9 and 3.2.10 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 3.2.1, we show this in columns (2), (3), (5), and (6).
- ✓ The standard errors ($Q_{seMage1}$ and $Q_{seMage2}$) of the unit changes at *Mage1* and *Mage2* are calculated using Equations 3.2.9 and 3.2.10. The standard errors are the absolute value of the product of the unit change (Q_{mage}), multiplied by the coefficient of variation (ES_{se} / ES) in the effect sizes at each age. In the example calculation in Exhibit 3.2.1, we show this in columns (3), (10), and (11).

- ✓ 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).
- ✓ 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).
- ✓ 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).
- ✓ 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).
- ✓ 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).
- ✓ 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. We previously implemented bounding rules on these dichotomous outcomes to prevent their draws from being below 0 or above 1. We have adjusted our methodology to account for the larger unit changes that would be possible if our base rate estimates are incorrect.

Exhibit 3.2.1

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	Compute the treatment group rate	Compute the unit change	Compute the percentage change	Distribute the percentage change to other years	Compute unit change	Distribute the coefficient of variation	Compute the standard error on the unit change
	Column									
	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
15			0.400				-0.185	-0.074	-0.500	0.037
16	-0.200	-0.500	0.420	0.342	-0.078	-0.185	-0.185	-0.078	-0.500	0.039
17			0.440				-0.159	-0.070	-0.500	0.035
18			0.460				-0.134	-0.061	-0.500	0.031
19			0.480				-0.108	-0.052	-0.500	0.026
20	-0.100	-1.500	0.500	0.459	-0.041	-0.082	-0.082	-0.041	-0.500	0.021
21			0.520				-0.082	-0.043	-1.500	0.064
22			0.540				-0.082	-0.044	-1.500	0.067
23			0.560				-0.082	-0.046	-1.500	0.069
24			0.580				-0.082	-0.048	-1.500	0.072
25			0.600				-0.082	-0.049	-1.500	0.074

Inputs	
15	Tage (age of person at time of treatment)
16	Mage1 (age of person when outcome first measured)
-0.200	ES1 (effect size at Mage1)
0.100	SE1 (Standard error at Mage1)
20	Mage2 (age of person when outcome is measured a second time)
-0.100	ES2 (effect size at Mage2)
0.150	SE2 (Standard error at Mage2)

3.3 Linked Effect Size Parameters

As noted in Section 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 the expression below, these linkages are the relationships between *Outcome*$_1$ and *Outcome*$_2$.

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

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 is estimated with the meta-analytic procedures described 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. Outcomes are calculated at an "age of link measurement" and take effect at an "age at which relationship begins."

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. The "age of link measurement" is also set as 18. This means that if a program has a direct impact on crime *after* age 18, then it is too late to activate these linked benefits of high school graduation.

In another example, preterm birth increases the likelihood of infant mortality, and thereby reduces the expected labor market earnings and other lifetime benefits for preterm infants compared to full-term infants. From a primary analysis of Washington State data (described in detail in WSIPP's Health Care Technical Appendix),[36] the effect size of preterm birth on infant mortality is 1.103 with a standard error of 0.072. Infant mortality by definition occurs within the first year of life, so we set the "age at which relationship begins" to 1 and present-value all future expected benefits back to age 1.

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 after the "age at which a relationship begins." When the link is calculated, we calculate the percentage change, which is distributed to all other ages, at the "age of link measurement." We list our current estimates for the linkages in Appendix I of this report.

3.4 Unit Changes from Linked Effect Sizes

For linkages between outcomes, the user enters a single effect size, standard error, the age at which to calculate the linked unit change, and the age at which to begin the measurement of the resulting unit change. To compute the linked unit change from these link effect sizes, we follow procedures analogous to those described in Section 3.2.

For continuous outcomes, as shown in Equation 3.4.1, 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:

$$(3.4.1)\ \ LinkQ_{LinkAge} = Base \quad \times ES_{link}$$

For dichotomous outcomes, as shown in Equation 3.4.2, 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).

$$(3.4.2)\ \ LinkQ_{LinkAge} = \left(\frac{(e^{ES_{link} \times 1.65} \times Base)}{(1 - Base + Base \times e^{ES_{link} \times 1.65})} - Base \right)$$

[36] Westley, E. & He, L. (2017). *Estimating effects of birth indicators on health care utilization costs and infant mortality: Technical appendix.* Olympia: Washington State Institute for Public Policy.

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 the following expression (described in Sections 2.1 and 3.3):

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

As the model runs, it searches for any possible links to the direct program outcomes measured and then implements the procedures in Sections 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 probability of recidivism by ten percentage points. This is the program unit change as described in Section 3.2 (*Program* on $Outcome_1$).

Further, suppose that a juvenile that engages in crime has a reduced probability of high school graduation of 20 percentage points. This is the linked unit change as described in Section 3.4 ($Outcome_1$ on $Outcome_2$).

Then, multiplying these two changes, FFT can be expected to lead to an increase in the high school graduation probability (*Program* on $Outcome_2$) of .02 (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 two 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 (see Section 3.2b). In the benefit-cost model, the benefits of FFT will then be computed for a 10 percentage point change in crime outcomes and a 2 percentage point change in high school graduation.

Again, these particular numbers are hypothetical and for illustrative purposes only; these numbers do not represent our actual current estimates for FFT.

Chapter 4: Procedures to Estimate the 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:

$$(4.0.1)\quad NPV_{tage} = \sum_{y=tage}^{N} \frac{Q_y \times P_y - C_y}{(1+Dis)^y}$$

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), where the treatment age plus the future years is equal to 100. 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 4.0.1 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 4.0.1, Q_y, is the estimated number of outcome "units" in year y produced by the program or policy. As shown in Equation 3.1.1, Q_y is dependent on the effect size and the base rate. Chapter 2 discussed the transformation of research literature into an effect size. Chapter 3 discussed the calculations used to go from an effect size to a unit change. This chapter will cover three different elements: the underlying framework applied to all outcome valuations, the Base Rate used in the calculation of quantity, and the value or price of that change in the quantity of an outcome, Py.

This chapter begins by discussing the background inputs to the benefit-cost model that affect the overall computation of NPV, then moves into the base rates and pricing of specific outcomes.

4.1 General Parameters

To make consistent comparisons, background assumptions are used to compute benefits and costs. These are discussed in this section.

4.1a 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 2018). When the model runs, all monetary values entered into the model are converted to the base year values with the price index (see Section 4.aaf).

4.1b 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 4.1.1 indicates that the net present value of a program, evaluated 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*.

$$(4.1.1)\quad NPV_{age} = \sum_{y=age}^{N} \frac{B_y - C_y}{(1+Dis)^y}$$

The model uses low, modal, and high discount rates in computations. When the model is run in non-simulation mode, the modal discount rate is used. In Monte Carlo simulation, each run randomly draws a discount rate from a triangular probability density distribution, with the low, modal, and high discount rates defining the triangle. Exhibit 4.1.1 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).[37] Similarly, the Congressional Budget Office has used a 3% real discount rate in its analyses of Social Security.[38] 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.[39] More recent work by Moore et al. (2013) restates the argument for using a 3.5% and 5% discount rate, while the Council of Economic Advisers (2017) has recommended a 2% discount rate.[40]

Exhibit 4.1.1
Discount Rates Used in Benefit-Cost Model

Range	Discount rate
Low value	0.020
Modal value	0.035
High value	0.050

4.1c Demographic Information

Several of the computations in the model require basic demographic information about the population in the jurisdiction to which the model is applied. Exhibit 4.1.2 displays a table with these inputs. For Washington State, we enter the current distribution of the total state population by single year of age 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 each year along with life expectancy. We use life table information produced by the U.S. Department of Health and Human Services Centers for Disease Control and Prevention.[41] 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 previous year's population to estimate the population for those ages.

4.1d 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 Section 4.5, 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.[42]

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.

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).[43] The VSL estimates are almost always much larger than the

[37] 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.
[38] Congressional Budget Office. (2012). *The 2012 long-term projections for social security: Additional information*. Washington, DC. Retrieved August 8, 2013.
[39] Heckman et al. (2010).
[40] Moore, M.A., Boardman, A.E., & Vining, A.R., (2013) More appropriate discounting: the rate of social time preference the value of the social discount rate. *Journal of Benefit-Cost Analysis*, 4(1), 1-16.
Council of Economic Advisers. (2017). Discounting for Public Policy: Theory and Recent Evidence on the Merits of Updating the Discount Rate. Council of Economic Advisers Issue Brief January 2017.
[41] Arias, E., Heron, M. & Xu, J. (2017). *United States life tables, 2014* (National Vital Statistics Reports vol. 66, no. 4). Washington, DC: United States Department of Health and Human Services, National Vital Statistics System, Table 1.
[42] 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.
[43] 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.

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 & Viscusi (2008), after constructing revealed preference wage equations, have provided recent estimates of VSL for ages 18 to 62.[44]

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 & Viscusi (2008) 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[45] and Social Security[46] if someone dies (See Exhibit 4.1.2). We also subtract the "human capital" derived benefits of changes to lifetime earnings (*LTE*), described elsewhere in this document. Thus, the general approach is given in the following equation:

$$(4.1.2) \quad VSLAdj = VSL - HC - SS - LTE$$

WSIPP's VSL model is driven with the parameters shown in Exhibit 4.1.3, along with the life table and public cost year information displayed in Exhibit 4.1.2. 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. (2010).[47] For the modal value, we compute the average between the high and low. These values are expressed in 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.

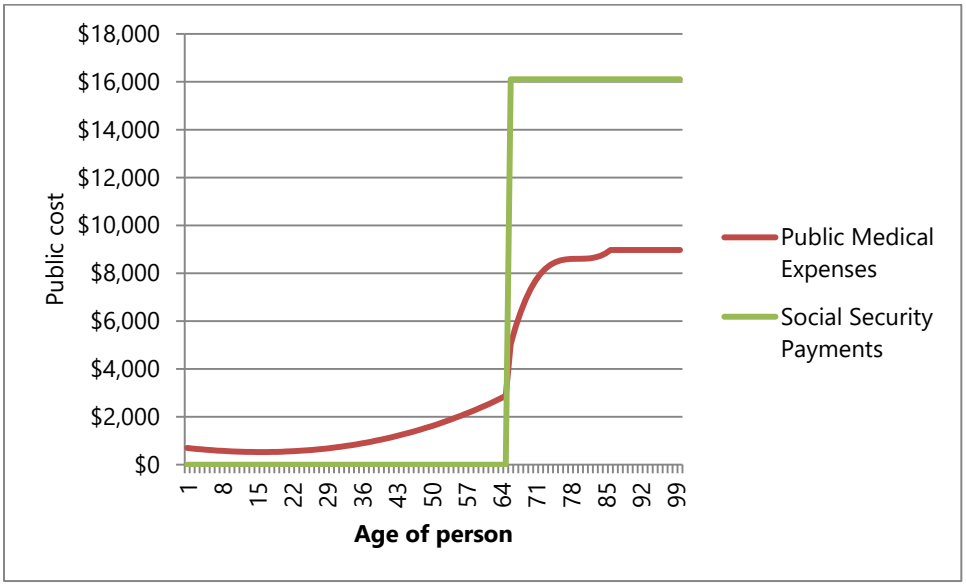
[44] Aldy, J.E., & Viscusi, W.K. (2008). Adjusting the value of a statistical life for age and cohort effects, *The Review of Economics and Statistics, 90*(3), 573-581.

[45] To estimate health care costs by age for the average person in the population, we used data from the 2015 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 the person receiving those services, was computed for 2015. To account for one-year spikes, third order polynomials were fit to the reported expenditures from ages 1 to 65 and from 65 to 100. Cost by year is displayed in Exhibit 4.1.2. These figures are adjusted for inflation and escalation in healthcare costs over time when applied.

[46] We use an average annual Social Security benefit of $14,657 in 2015 dollars, from age 66 on (Monthly Statistical Snapshot, June 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.

[47] 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 4.1.2
Value of a Public Cost Year



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 4.1.3 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 as described in XXX.

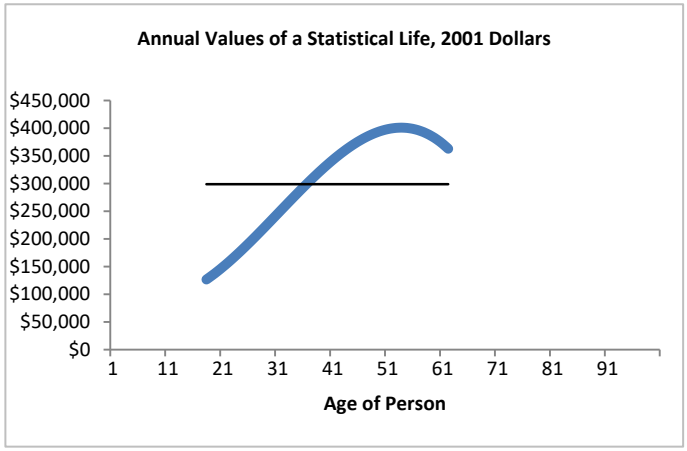
$$(4.1.3)\ \ VSLY = \frac{disrate \times VSL}{1 - (1 + disrate)^{-L}}$$

Exhibit 4.1.3
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^^{2}	0.65
Age^^{3}	-0.007
Post-age 62 exponential change rate	0.00
Pre-age 18 multiplier	1.0

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 4.1.3 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.

Exhibit 4.1.4



The Aldy & 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 & Hersch (2008) [48] 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 4.1.3 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:

$$(4.1.4) \quad VSLY_y = VSLY_{62} \times (1 + esc)^{(y-62+1)}$$

Valuation of Reductions in Infant Mortality Risk. Some studies directly measure the likelihood of mortality in the year following birth. Additionally, WSIPP has estimated causal links between other birth outcomes (such as low birthweight or preterm births) and increased mortality risk in the year following birth.[49] For direct and indirect valuation of infant mortality, we use the same procedures described above to value the statistical life years foregone when mortality risk increases. For ages less than 18 (the earliest age for which a *VSLY* can be estimated with the Kniesner and Viscusi data), our review of the evidence did not reveal a consensus around valuing a statistical life-year for youth. Although one study, Hammitt & Haninger (2010), found through stated preference methodology that young children's lives are valued higher than adult lives, we take a cautious approach and set the value of a statistical life year for ages less than 18 equal to that of the 18th year of life.[50]

4.1e 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.[51] 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 x 50% - $1,000 x 50%). In the actual run of the model, these calculations are carried out for each year of cash flows.

$$(4.1.5) \quad DWL_{age} = \sum_{y=age}^{N} \frac{(B_y - C_y) \times DWL\%}{(1 + Dis)^y}$$

[48] Viscusi, W.K., & Hersch, J. (2008). The mortality cost to smokers. *Journal of Health Economics, 27*(4), 943-958.
[49] Westley & He (2017).
[50] We had previously used the ratio of the VSL for children relative to adults (1.7) reported by 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.
[51] 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.

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 program.[52] Also following Heckman et al. (2010), we do not apply any deadweight cost calculations to estimated taxes obtained from earnings-related outcomes.[53]

4.1f 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. The general inflation index used by WSIPP is the Implicit Price Deflator for Personal Consumption Expenditures produced by the Bureau of Economic Analysis.[54] 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.[55] We use the BEA Implicit Price Deflator for Personal Consumption Expenditures for Health Services.

4.1g Tax Rates

The benefit-cost model uses average tax rates for several calculations. We used the aggregate total from the Tax Foundation from 2016 to represent a combination of all kinds of taxes paid (income, sales, property, and other), as a percentage of income.[56] This value and the breakdown are displayed in Exhibit 4.1.5.

Exhibit 4.1.5
Tax Rates

	Percent of total, by source		
Total tax rate	Federal	State	Local
0.2986	0.6413	0.2027	0.1560

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.[57]

4.1h 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 the fiscal staff of the Washington State Legislature.

[52] Heckman et al. (2010).
[53] Ibid, see section J of the Heckman Appendix.
[54] 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 August 14th, 2019.
[55] 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 August 14th, 2019.
[56] We looked at data from two separate sources: York et al. (2019). *Tax Freedom Day® 2019 is April 16th*. Washington, DC: Tax Foundation, Retrieved August 14, 2019, and Citizens for Tax Justice (2016). Who pays taxes in America in 2016? Washington, DC. Retrieved April 13, 2017. The first source gave a federal estimate of a total effective tax rate of 29.8%, while the second source gave an estimate of 29.9%. Because these numbers were so similar, we used the Tax Foundation number of 29.8%.
[57] 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). The Tax Policy Center performs calculations on Bureau of Economic Analysis Tables 3.2, 3.20, and 3.21. The method was retrieved May 18, 2016. Bureau of Economic Analysis tables can be found at. U.S. Bureau of Economic Analysis, *Section 3 – Government Current Receipts and Expenditures*. Retrieved August 14th, 2019.

4.2 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.[58] Section 4.2a discusses the common data sources we use for all of the estimates involving labor market earnings, including those using a human capital approach as well as those derived from directly measured employment and earnings outcomes. 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 (Section 4.8)
- ✓ Standardized student test scores (Section 4.8)
- ✓ Higher education achievement (Section 4.8)
- ✓ Morbidity and mortality costs of alcohol and illicit drug disorders, and regular smoking (Section 4.5)
- ✓ Morbidity and mortality costs of mental health disorders (Section 4.6)
- ✓ Morbidity and mortality costs of diabetes and obesity (Section 4.7b)
- ✓ Morbidity and mortality costs of child abuse and neglect (Section 4.10)
- ✓ Earnings (Section 4.2)
- ✓ Employment (Section 4.2)

When we monetize specific programs, we make an effort to match the expected earnings of that population. One way the model organizes earnings is by educational subgroup. These educational subgroup calculations are described in Section 4.2b.

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 4.2c. Calculations of variations in labor market earnings and employment by various health conditions, mental health disorders, and substance use disorders are described in Section 4.2d. Outcomes may directly change earnings or change earnings through the probability of employment. These calculations are described in Section 4.2e. Finally, we discuss our method for calculating Public assistance and food assistance costs in Section 4.2f.

4.2a 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 U.S. Census Bureau's March Supplement to the Current Population Survey (CPS), which annually provides cross-sectional data for earnings by age and by educational status.[59] To keep the model as simple as possible, we gather "person variables" from the CPS summary files, including 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 U.S. population, not just those living in Washington State.

To prevent our long-term earnings projections from being based on a single year of data, we compute the average employment rates and present-valued earnings across an entire "trough-to-trough" business cycle. This allows us to avoid potential bias from single-year earnings and employment data that may be particularly strong or weak.

We use data that attempts to match the November 2001 to June 2009 business cycle as reported by the National Bureau of Economic Research (NBER).[60] Thus, we use the 2002 through 2010 March CPS files given that these files cover earnings for

[58] 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.

[59] The data are accessed from the "DataFerrett" application of the US Department of Commerce, Bureau of the Census.

[60] 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.

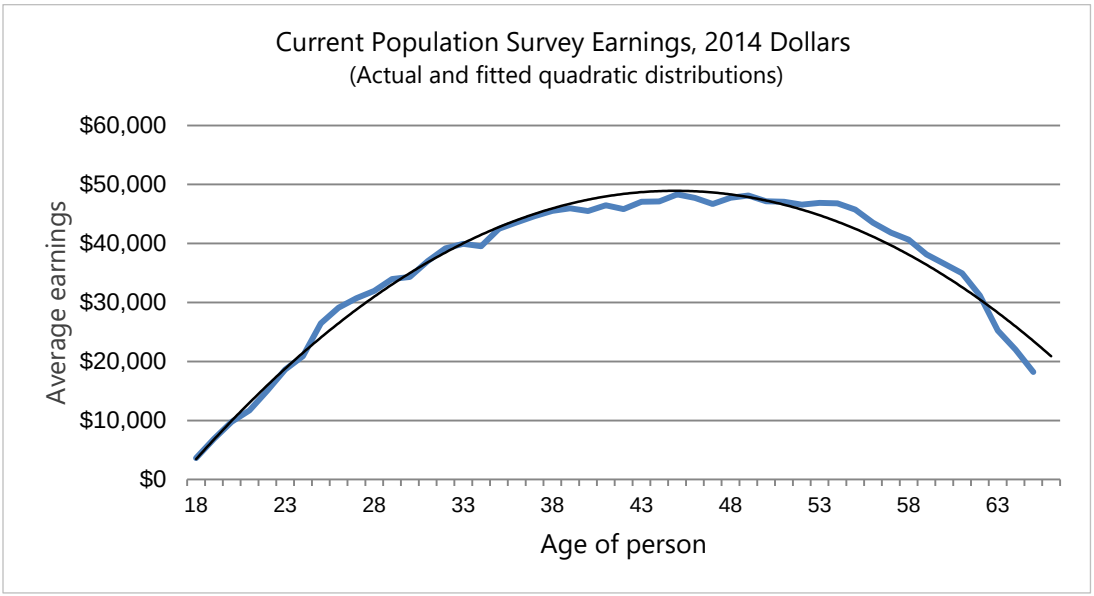
the prior year. The sample was restricted to persons aged 18 to 65 inclusive. It was weighted by the CPS March supplement final weight scaled such that the sum of the weights is equal to the number of unweighted observations in the data. From this sample, we ran a regression to compute average earnings per person by single year of age. We refer to this as *EarnAll*.

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

$$(4.2.1) \quad EarnAll_y = \beta 0 + \beta 1 * \mathrm{AGE} + \beta 2 * \mathrm{AGE}^2 + \beta 3 * \mathrm{YR}_{2003} + \beta 4 * \mathrm{YR}_{2004} + \ldots + \beta 10 * \mathrm{YR}_{2010}$$

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 measures 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 Implicit Price Deflator (IPD) described in more detail in Section 4.aaf. Further adjustments, described below, adjust the data to match the future labor market in Washington.

Exhibit 4.2.1



State-Specific Adjustment for Wages. We use an adjustment ratio to approximate earnings in Washington State relative to the national average. The CPS sampling was not designed to be representative at the state level, so we use information from the 1-year American Community Survey Public Use Microdata Sample (PUMS) for the years 2001 to 2009 to match the business cycle used in our general earnings calculations from the CPS.[61] We estimate a similar equation as that on earnings level but include a Washington State dummy variable. We divide the predicted earnings including the Washington State dummy variable by the observed earning in the whole country.[62] That percentage differential in earnings is used to adjust the national earnings calculated by the CPS to Washington.

[61] Datafiles are downloaded from the US Department of Commerce, Bureau of the Census.
[62] The variables in the regression included age, age^2, a WA state dummy and year dummies. In the PUMS, earnings is the sum of two variables: wage and salary earnings (WAGP) and self-employment earnings (SEMP).

Growth Rates in Earnings. Since these CPS data are cross sections for the most recent CPS year, and since our benefit-cost analysis reflects lifecycle 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).[63] We adjust the series for inflation using the U.S. Implicit Price Deflator for Personal Consumption Expenditures from the U.S. Department of Commerce (see Section 4.aaf). 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 wages 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).[64] The ECEC includes paid leave, supplemental pay, insurance, retirement and savings, and legally required benefits.[65]

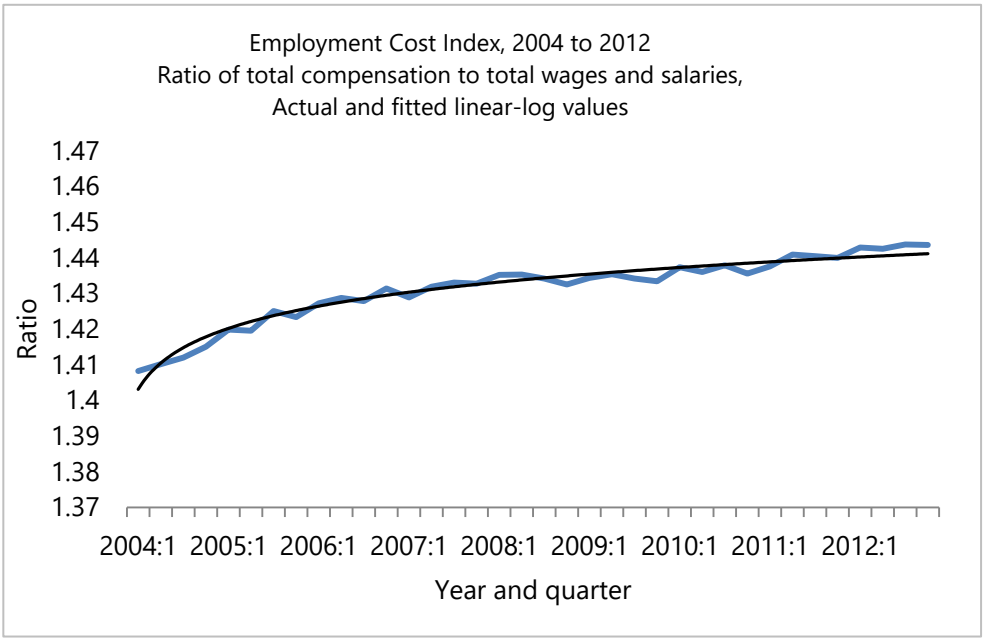
Exhibit 4.2.2

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

Exhibit 4.2.3 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 4.2.3



[63] We use a sample including persons ages 18-65 for our calculations of the adjustment of Washington State-specific wages and the growth in earnings.

[64] U.S. Bureau of Labor Statistics. (2016). *Employer costs for employee compensation—December 2015* (USDL-16-0463), Washington DC. Data retrieved March 30, 2016.

[65] Ibid.

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 the 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.aac, 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:

$$(4.2.2) \quad ModEarnAll_y = \left(\left(EarnAll_y \times (1 + EscAll)^{y-tage}\right) \times (FAll \times (1 + EscFAll)^{y-tage}) \times \left(IPD_{base}/IPD_{cps}\right) \times StateAdjAll\right) \times ProbLife_y$$

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, multiplied by 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.

This same process is used to model earnings for the subpopulations described below.

4.2b 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 4.8c and Section 4.8b). 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 4.2a using subsets of the data sample for four educational status groupings (and two subset groupings):

- ✓ Those who did not report completing high school but completed 7th grade or higher
- ✓ Those who reported completing high school with a diploma)
- ✓ Those with some college but no 4-year degree
- ✓ Those with some college but no degree of any type
- ✓ Those with a 2-year degree
- ✓ Those with a 4-year degree or more

For each of these six groups, we replicate the regressions and modeling to determine separate earnings by age distributions and different earnings growth parameters, displayed in Exhibits 4.2.4 and Exhibit 4.2.5.[66] We assume that students do not earn money for the time spent in higher education, and so for college populations, we set earnings to zero for the expected time spent in college (described in Section 4.8b).

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.

[66] The CPS does not ask about associate's degrees before 1992. To better match our business cycle approach to growth rates in earnings, we use the long term growth rate in earnings for the some college population for the two some college subset populations.

Exhibit 4.2.4

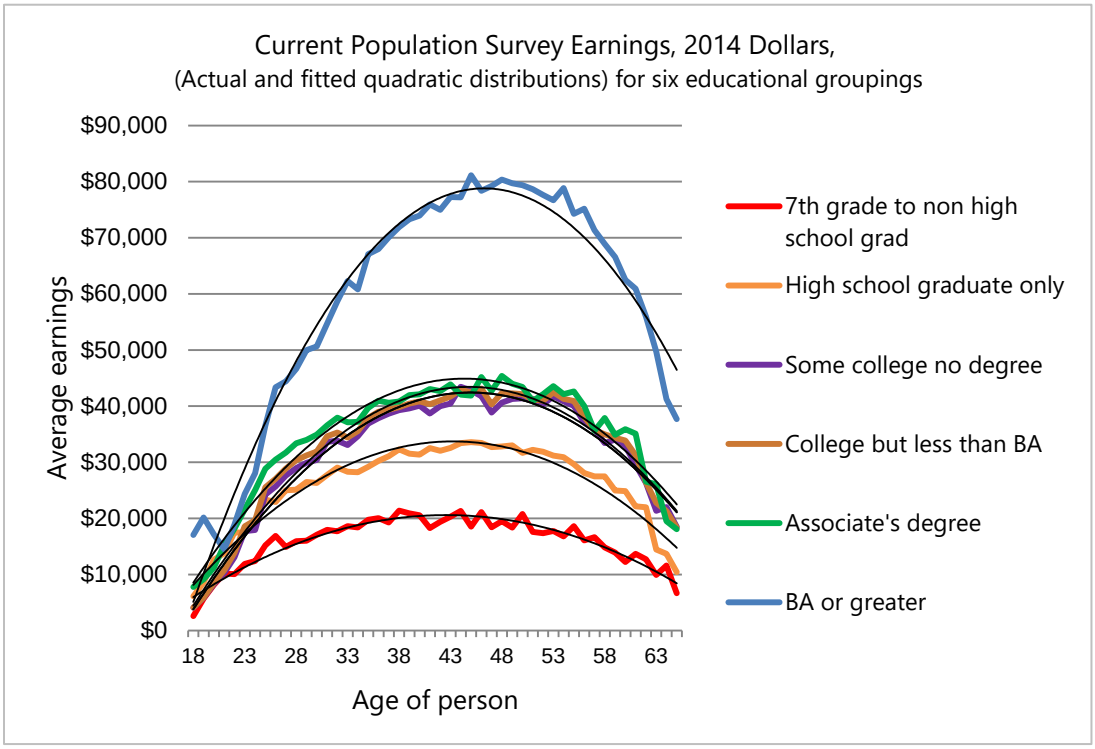
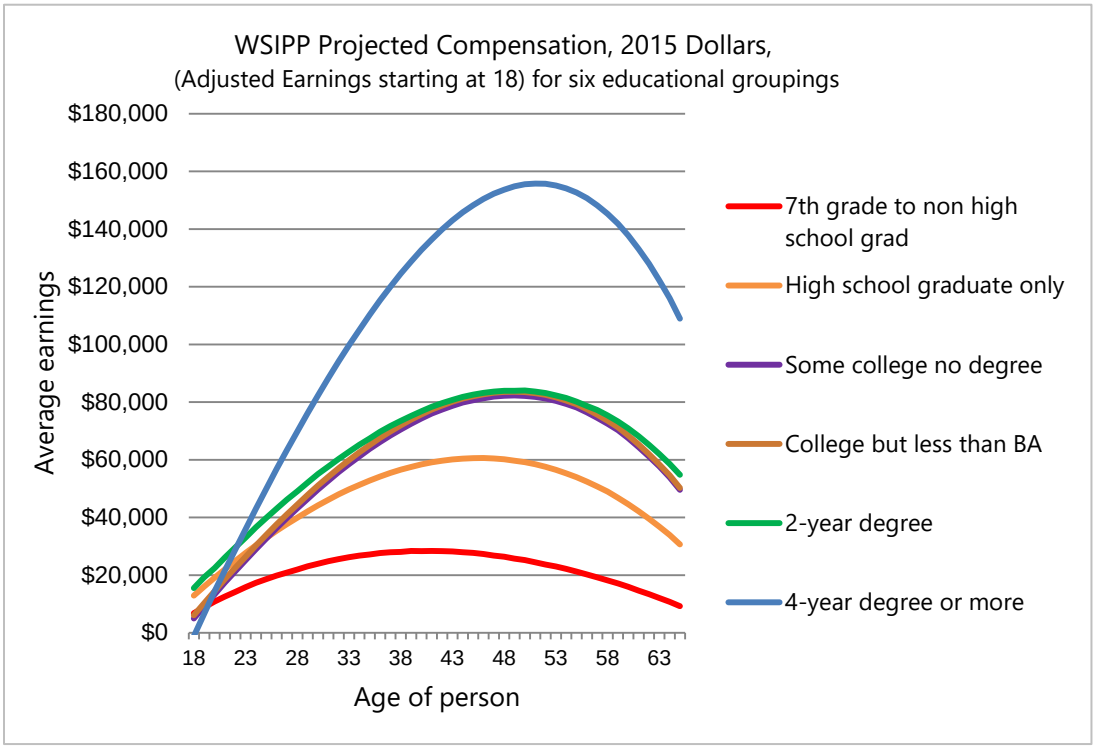


Exhibit 4.2.5
Earnings Adjustment Parameters by Educational Attainment

	7th grade to non-high school	High school graduate only	Some college, no degree of any type	College but less than 4-year degree	2-year degree	4-year degree or more
Annual real growth rates in earnings	-0.0062	0.0053	0.0095	0.0095	0.0095	0.0115
Benefits-to-earnings ratio	1.441	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	0.00041
Ratio of state to national earnings	1.079	1.074	1.007	1.003	0.986	0.935

These adjustment parameters are applied as described in Equation 4.2.2. Exhibit 4.2.6 below displays the 2015 projected earnings for a program that begins in year 18.

Exhibit 4.2.6



4.2c Earnings by Other Population Characteristics

The WSIPP model also values earnings for certain policy-relevant sub-populations. For example, WSIPP estimates values for some programs that directly target the labor market. We, therefore, segment 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 4.2a 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, and
- ✓ Not employed single parents (high school 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 percentage of the subgroup that is employed (has earnings greater than zero). We calculate growth parameters and state adjustment factors based on combinations of relevant education subgroups. Our factors are displayed in Exhibits 4.2.7 and Exhibit 4.2.8.

WSIPP also projects expected earnings for two additional groups: individuals with a serious mental illness and individuals previously involved in the criminal justice system. For each of these populations, we project earnings by multiplying our modified earnings for all people by an adjustment factor as listed in Exhibit 4.2.9.

Exhibit 4.2.7

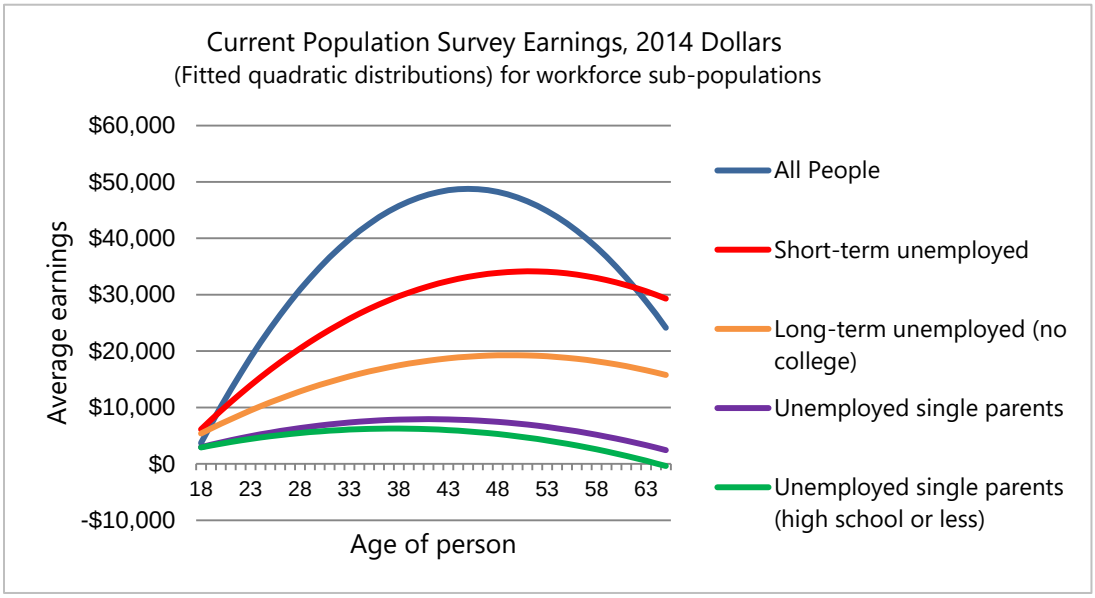


Exhibit 4.2.8

Earnings Adjustment Parameters by Workforce Population

	All people	Short-term unemployed^	Long-term unemployed (no college)*	Unemployed single parents^	Unemployed single parents (high school or less)#
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

Notes:
^ Subset of all people.
* Average of factors for less than high school, high school graduate, and some college education subgroups.
# Average of factors for less than high school and high school graduate education subgroups.

Exhibit 4.2.9

Earnings Adjustment Parameters by Educational Attainment

Population	Ratio of earnings for subgroup to all people	Probability of Employment
Serious mental illness	0.220^	0.334
Previous criminal justice system involvement	0.359*	

Notes:

^ This factor was estimated by comparing the average monthly earnings of Washington State Department of Social and Health Services (DSHS) clients with serious mental illness[67] with the average earnings of all workers from the CPS. This factor forms the variable *PctSMIEarn* used in Equation 4.6.5.

This number is the percent of DSHS clients considered to be seriously mentally ill who have any employment.[68]

* Number represents the ratio of the average earnings for the DSHS criminally involved population compared to the general population.[69] This factor is used to compute the base level of earnings when monetizing earnings for Adult Criminal Justice programs which measure earnings.

4.2d 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 focus 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 4.2.9).). We do this across a broad age range (18-65) as well as for a more limited population of older adults (50-65) who match the age range of certain targeted programs. As mentioned in Section 4.2d, to prevent our long term earnings projections from being based on a single year of data, we compute the average employment rates and present-valued earnings across an entire "trough-to-trough" business cycle. This allows us to avoid potential bias from single-year earnings and employment data that may be particularly strong or weak. We then apply the effect of the condition or disorder on the 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 4.5d.

Exhibit 4.2.10

Base Assumptions for Earnings and Employment, Business Cycle
Developed from 2002-2010 March Supplement of the CPS (2014 dollars)

	Mean earnings of workers	SD of earnings of workers	Percent of population that works
Ages 18-65	47,075	56,025	78.04%
Ages 50-65	56,433	67,018	70.67%

[67] Average annual wages for calendar year 2015 ($10,435) provided by D. Mancuso, Director, DSHS Research and Data Analysis Division (personal communication, April 3, 2017).

[68] Ibid.

[69] Average annual earnings for workers with previous arrest and booking for calendar year 2017 ($12,088) provided by J. Mayfield, DSHS Research and Data Analysis Division (personal communication, October 9, 2018).

4.2e Valuation of Earnings and Employment Outcomes

This section describes WSIPP's benefit-cost modeling of labor market outcomes that are measured directly in program evaluations, and 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 percentage 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. Estimated program effects on earnings are calculated with a meta-analysis of elasticity "effect sizes" which results in an expected percentage change in earnings. We multiply this estimated percentage change in earnings by the projected earnings for the specified population in each year (see Section 4.2c for a description of these populations). After adjusting for the loss of earnings due to death in the participating population, the percentage change is applied to the projected stream of annual earnings for the specified population produced by Equation 4.2.2.

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:

$$(4.2.3) \quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{(\Delta Emp_y \times PopEarn)}{(1+Dis)^{(y-tage+1)}}$$

PopEarn is estimated by dividing the expected earnings of the population analyzed by the percentage 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.

4.2f Valuation of Public Assistance Outcomes

Separately from measures of labor market earnings, we estimate program effects on government financial assistance. A portion of public assistance costs is treated as a transfer payment 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.

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.[70] 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 4.2.11. To estimate the proportion of total TANF/SFA expenditures that come from state versus federal sources, we use data reported by the TANF program.

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 directly provide other, non-cash-assistance services, any additional costs of these programs are the costs to administer the program.

[70] Advice on categories to exclude (expenditures that would not be expected to be reduced if the adult caseload reduced) was provided by S. Ebben, Economic Services Administration (personal communication, August 28, 2015).

Exhibit 4.2.1 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.

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. In terms of the timing of these expected benefits, we estimate that they occur for some duration between the age of treatment and the age of measurement. Thus, the total estimated increase in assistance is evenly divided among all years between the age of treatment and the age at first measurement.

Exhibit 4.2.11
Public Assistance Parameters

	Cash assistance	Food assistance
Average monthly benefit	$407.80[1]	$215.57[2]
Administrative proportion	1.74[3]	0.13[4]
Average months on assistance	12.7[5]	40.5[6]
SD of months on assistance	12.2[5]	36.8[6]
Age at which assistance begins	18	18
Year of dollars	2014	2014
Proportion from state sources	0.35[7]	0.07[8]
Proportion from local sources	0.00[7]	0.00[8]
Proportion from federal sources	0.65[7]	0.93[8]

Notes:
[1] Average monthly payment per case for FY2018. Source: 2018 TANF Work First as of September 2019.
[2] Average monthly payment per case for FY2018. Source: 2018 Basic Food as of September 2019.
[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 FY2018. 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 2016.
[5] Total length of assistance and standard deviation in months computed using a cohort of adult clients entering TANF/SFA in January 2014 and following them through December 2018. Source: ESA-EMAPS Report #4786 using the ACES Data Warehouse as of July 2019.
[6] Total length of assistance and standard deviation in months computed using a cohort of adult clients entering SNAP/FAP in January 2014 and following them through December 2018. Source: ESA-EMAPS Report #4786 using the ACES Data Warehouse as of July 2019.
[7] Proportion of costs borne by state and federal sources are derived from assistance and non-assistance categories reported in TANF Financial Data for FY2018, 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 2016. 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.

4.3 Valuation of Health Care Outcomes

The purpose of WSIPP's health model is to inform the Washington State Legislature whether there are economically attractive evidence-based policies that, if implemented well, can achieve reductions in the cost of care and/or improvements in health conditions. WSIPP's health model monetizes the projected lifecycle costs and benefits of programs or policies that have been shown to achieve improvements—today and in the future—in 1) health care costs and resource utilization; 2) health outcomes; and 3) 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.

We describe general parameters and the data sources which we use when calculating health care costs throughout the benefit-cost model in Section 4.3a. The model estimates the value of changes in health care costs and health care resource utilization for the specific populations targeted by the interventions we have investigated so far, such as chronically ill individuals or new mothers receiving Medicaid. In addition to the total costs of health care for individuals, the utilization measures include hospitalization (both general and psychiatric), hospital readmissions, and emergency room visits. We discuss the valuation of changes in health care costs and resource utilization in Section 4.3b.

WSIPP's model also monetizes certain health-related outcomes, including falls among older adults; the cost of cesarean sections for mothers; and the costs of preterm, low-birthweight, and neonatal intensive care unit (NICU) admissions; and births that are small for gestational age, for both mothers and infants. We discuss the valuation of an average fall in Section 4.3c and the valuation of maternal and infant health outcomes in Section 4.3d.

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 outcomes considered in this model. Depending on each particular outcome, the following benefit or cost categories are included in WSIPP's model:

- ✓ Total costs of care, to the degree that interventions (e.g., patient-centered medical homes) reduce costs.
- ✓ Hospital admission, readmission, and emergency department costs, to the degree that interventions (e.g., case management for frequent ED user, care transition programs) reduce utilization.
- ✓ Hospital costs in the first year after birth for mothers and infants stemming from birth outcomes (i.e., preterm birth, low- and very low-birthweight births, small for gestational age infants, admissions to NICU facilities), to the degree that interventions (e.g., smoking cessation for pregnant women) reduce poor outcomes.
- ✓ Hospital costs in the first year after a fall for older adults, to the degree that interventions (e.g., exercise programs for fall prevention) reduce the incidence rate of falls.
- ✓ Total costs of cesarean sections for mothers, to the degree that interventions can reduce unnecessary c-section rates.
- ✓ Falls and infant mortality Value of a statistical life (VSL) estimates, net of labor market gains, applied to the change in mortality estimated to be caused by, along with lifetime earnings lost because of premature death (mortality) caused by health conditions.

4.3a General Health Care Parameters

Total personal health care expenditures are collected from the Centers for Medicare & Medicaid Services at the U.S. Department of Health & Human Services. We use the ratio between pharmaceutical/drug expenditures and inpatient hospital expenditures to compute an added drug cost for every hospital visit we monetize throughout the model. A hospital cost-to-charge ratio for Washington State is computed with 2011 data from the Healthcare Cost and Utilization Project (HCUP) of the U.S. Department of Health & Human Services. Total annual emergency room visits in Washington for 2008 are computed from data compiled by the Washington State Hospital Association.

Exhibit 4.3.1

General Health Care Parameters

Parameter	Value
Total National personal health care expenditures^	$2,834,000,000,000
Hospital care	$1,082,500,000,000
Drugs	$328,600,000,000
Hospital cost-to-charge ratio#	0.346
Emergency department cost-to-charge ratio*	1.0
Emergency department admissions, 2008^^	1,997,069

Notes:

^ Centers for Medicare & Medicaid Services. National Health Expenditure Tables—Table 2. Retrieved November 16, 2018.

Agency for Healthcare Research and Quality, Healthcare Cost and Utilization Project.

* WSIPP assumption.

^^ Number calculated from a number of ED visits per 1,000 people in Washington from Kaiser Family Foundation. Data are for community hospitals. Data retrieved from Kaiser Family Foundation Website October 2018.

One of the datasets we use to estimate health care costs is MEPS, a nationally representative large-scale survey of American families, medical providers, and employers who report on health care 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 the quality of life. 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. 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 providers (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, and ambulance. There are some limitations to using MEPS data, including that negotiated health prices may not reflect the true cost of care, and MEPS data do not include uncompensated care. We typically perform calculations using survey weights.

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 4.3.2 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, U.S. Department of Health & Human Services.[71] 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.

[71] Centers for Medicare & Medicaid Services. (n.d.). *National health expenditure projections 2009-2019.* United States Department of Health & Human Services. Retrieved June 30, 2011.

Exhibit 4.3.2
Proportion of Health Care Costs by Source

	Total cost by perspective			Taxpayer cost by payer		
	Participant	Taxpayer	Other	State	Local	Federal
General health care^	12.21%	43.20%	44.58%	14.72%	0.00%	85.28%
Emergency department^	9.9%	36.45%	53.65%	18.19%	0.00%	81.81%
Mental health costs*	1.10%	80.80%	18.20%	27.26%	0.00%	72.74%
ATOD treatment#	12.71%	38.97%	48.32%	45.79%	3.69%	50.51%
General hospital^	2.12%	49.29%	48.59%	10.64%	0.00%	86.14%
Drug/pharmacy^	21.80%	44.90%	33.30%	15.65%	0.00%	84.35%

Notes:
^ WSIPP calculation from 2015 Medical Expenditure Panel Survey (MEPS) data for all ages.
* 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.
# ATOD = alcohol, tobacco, and other drugs. 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."*

4.3b Valuing Measured Changes in Health Care Costs and Resource Utilization

We monetize differences in health care expenditures in two different ways. The first is when studies measure changes to total health care costs. The second applies to direct measures of utilization of various components of the health care system (e.g., hospitals, emergency departments).

Changes to Total Health Care Expenditures. Some studies look at the effect of programs in terms of the % change in overall healthcare spending. The benefit-cost model directly monetizes these changes to total health care expenditures. The percent change in health care costs as a result of participation in a program is multiplied by the average annual cost for health care for the specified population. Typically, program evaluations only report changes in health care costs over a brief follow-up period (i.e., three years or less). Therefore, we only model these changes in costs for the reported period.

Exhibit 4.3.3
Total Health Care Cost Parameters

	Chronically ill adults	General population
Average annual cost for health care^	$12,848	$4,978
Standard deviation on cost	$23,666	$15,132
Year of dollars	2015	2015

Notes:
Chronically ill adults are those who are at least 18 years old and have been diagnosed with one or more of the following conditions: coronary heart disease, angina, heart attack, other heart diseases, diabetes, stroke, or emphysema.
^ WSIPP calculation from 2015 Medical Expenditure Panel Survey data (MEPS).

Health Care Resource Utilization. Second, we describe the parameters for estimating the benefits of program-related changes in specific health care resource utilization (see Exhibits 4.3.2-4.3.6). WSIPP monetizes measured increases in hospitalization, psychiatric hospitalization, emergency department use, and hospital readmissions. To model the monetary benefits of changing the utilization of these health care resources, we multiply the average cost of the measured resource for the specified population by the unit change produced from the program effect size and the base rate for that population. For programs with measures of multiple resources, we sum the changes into a single measure of service utilization. For most resources, the effects produced by programs are time-limited, e.g., reducing the likelihood of a hospitalization produces monetary benefits for a single year. The value of changes to health care resource utilization is represented by the following equation:

$$(4.3.1)\ \Delta HealthCareCosts_{type} = \frac{\Delta Use_{type} \times HeathCareCost_{type} \times (1 + HCCEsc)^{y-tage}}{(1 + Dis)^{y-tage}}$$

For each health care resource *type* measured by studies of a program, we multiply the unit change by the annual cost of that resource for the specified population adjusted for escalation and discounted to the year of treatment.

Exhibit 4.3.4
Hospitalization Parameters

	Children with asthma	Frequent emergency department users#	General population
Average cost for a hospitalization^	$6,202	$36,714	$20,811
Standard deviation on cost	$8,224	$40,446	$33,384
Year of dollars	2015	2015	2015
Annual likelihood of hospital admission*	2.34%	64.22%	6.42%

Notes:
Hospitalization parameters for older adults hospitalized due to a fall are described in section 4.3.c
Frequent emergency department users are adults who visited the ED five times or more within a single year.
^ WSIPP calculation from 2015 Medical Expenditure Panel Survey (MEPS) data.
* Of those in population, the proportion who were admitted to the hospital in a single year (MEPS).

Exhibit 4.3.5
Hospital Readmission Parameters

	Chronically ill adults	General population
Average cost for a readmission^	$20,166	$18,043
Standard deviation on cost	$31,808	$25,717
Year of dollars	2012	2012
Likelihood of readmission within 30 days after discharge*	24.8%	9.1%

Notes:
Chronically ill adults are those who are at least 45 years old and have been diagnosed with one or more of the following conditions: coronary heart disease, angina, heart attack, other heart diseases, diabetes, stroke, or emphysema.
^ WSIPP calculation from 2012 Medical Expenditure Panel Survey (MEPS) data. Unlike other calculations in this section, these numbers were not calculated with survey weights.
* 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).
Of those in population and who had had at least one admission to the hospital, the proportion who were re-admitted to the hospital within 30 days of discharge (MEPS).

Exhibit 4.3.6
Emergency Department Parameters

	Children with asthma	Frequent emergency department users	General population
Average cost for an ED visit^	$787	$6,803	$1,555
Standard deviation on cost	$1,388	$7,886	$3,587
Year of dollars	2015	2015	2015
Annual likelihood of ED visit*	18.03%	50.00%^^	14.22%

Notes:
Frequent emergency department users are adults who visited the ED five times or more within a single year.
^ WSIPP calculation from 2015 Medical Expenditure Panel Survey data (MEPS) for those with an ED visit.
* Of those in population, proportion who visited the emergency department in a single year (MEPS).
^^ 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.

4.3c Valuing Falls for Older Adults

In this section, we describe our method for valuing a fall in the older adult population. The Centers for Disease Control and Prevention (CDC) estimate that 28.7% of older adults reported falling in 2014.[72] Falls vary in levels of severity; while some falls do not require medical attention, others can result in serious injury or death. We calculate the expected number of falls per person per year and the probability that any individual fall will result in hospitalization or death.

Fall Incidence. WSIPP uses an incidence rate of falls calculated from by the Washington State sample of the Behavioral Risk Factor Surveillance Survey (BRFSS), a national survey designed to provide valid state-level information about behavioral risk factors and health. We used responses to the question, "In the past 12 months, how many times have you fallen?"[73] We use the BRFSS CDC weighted n's of respondents by age group to compute a weighted average of the number of falls.[74] Because individuals who died as a result of a fall are not present in the survey sample, we add the number of deaths due to falls to both sides of the fall rate. For each age group (age 65–69, 70–74, 75–79, and 80+) we compute an average incidence rate

[72] Bergen, G., Stevens, M.R., & Burns, E,R. Falls and fall injuries among adults aged≥ 65 years—United States, 2014. *Morbidity and mortality weekly report, 65*.
[73] Centers for Disease Control. (n.d.). *Behavioral risk factor surveillance system 2014 codebook report.* Retrieved September 2017.
[74] In reported BRFSS data, falls are top-coded—that is, falls are reported as categorical outcomes, with classifications of 0, 1, 2, 3, 4, and > 5 falls. This data limitation likely lowers the expected fall rate, as individuals who fall greater than five times in a year are coded as having reported falls—regardless of the actual number of falls. While the CDC does a calculation of uncensored BRFSS data, it is not available by age group. This uncensored fall number captures the falls of chronic fallers, including those with co-occurring risk factors. We have chosen to use the censored fall rate, which allows for age-group specific rates and avoids overweighting chronic fallers. A comparison of the censored and uncensored rates indicates that the resulting estimate may be missing up to 25% of all falls.

of falls over the three most recent years of BRFSS surveys in which falls questions were asked.[75] We also calculate a fall incidence rate for those with a high risk of experiencing a fall due to the presence of falls risk factors apart from age. A meta-analysis of falls risk factors by Deandrea et al. (2010) estimated the increased risk of falling for community-dwelling older adults with particular risk factors for falls.[76] This study estimated that, on average, individuals with a previous history of falls have 2.77 times greater odds of experiencing a fall than older adults without a previous history of falls. We use this estimated odds ratio to calculate our fall incidence rate for a high-risk population for each age group. The average number of falls by age group and population shown in Exhibit 4.3.7 is the base incidence rate of falls in our model.

Exhibit 4.3.7
Fall Rates

Age group	Fall incidence rate (falls per person per year)	
	General population	High-risk population
65–69	0.608	1.684
70–74	0.631	1.747
75–79	0.613	1.699
80+	0.690	1.911

Each fall results in some chance of hospitalization and some chance of death.[77] Our model accounts for fall-related hospitalization and fall-related death because these secondary outcomes have related costs. We estimate the likelihood that a fall results in hospitalization or death using information from the Washington State Department of Health's Community Health Assessment Tool (CHAT), a state data system for population-level data sets.[78] These data include fall hospitalization and death rates as well as population estimates in Washington in five-year age groups for the years 2012 and 2014.[79] We calculate a rate of hospitalizations due to falls as the number of hospitalizations over the number of falls in each age group. We repeat this process with the number of deaths to calculate the rate of death from falls for each age group. Exhibit 4.3.8 shows the expected number of falls as well as the percent of falls that result in a hospitalization and the percent of falls that result in death.

Mortality Attributable to Falls. We estimate the likelihood that a fall will result in death. The chance of death attributable to a fall is related to the age of the individuals who falls, as detailed in Exhibit 4.3.8. WSIPP's model values mortality using our value of statistical life (VSL) method described in Section 4.1d. Since our model values death as VSL rather than through costs associated with the death itself, we are not double-counting when a fall results in both a hospitalization and a death.

Exhibit 4.3.8
Likelihood of Hospitalization or Death After a Fall

Age group	Likelihood of hospitalization	Likelihood of death
65–69	0.88%	0.02%
70–74	1.41%	0.04%
75–79	2.43%	0.08%
80+	5.58%	0.36%

[75] The question appears in the surveys of even-numbered years. We use information from 2012, 2014 and 2016. WSIPP collected the incidence rate from the BRFSS WEAT online data system for survey years 2012 and 2014. Centers for Disease Control and Prevention (n.d). *BRFSS Data Access from the WEAT Tool*. Retrieved September 2017. Information from the 2016 BRFSS was provided by personal communication with Mark Serafin of the Washington State Department of Health on October 20, 2017.
[76] Deandrea, S., Lucenteforte, E., Bravi, F., Foschi, R., La, V.C., & Negri, E. (2010). Risk factors for falls in community-dwelling older people: A systematic review and meta-analysis. *Epidemiology*, *21*(5), 658-668.
[77] WSIPP assumes that the probability of hospitalization or death is constant across each individual fall.
[78] CHAT death rates are based on Washington death certificate data while hospitalization information is based on hospital discharge data from Washington (CHARS) and Oregon.
[79] Carolyn Ham, Washington State Department of Health (personal communication, November 8, 2017).

Health Care Costs Attributable to Falls. WSIPP reviewed the literature on falls among older adults to determine the average health care costs incurred for a fall. In our review, we found varying estimates across sources. We prioritize cost estimates that come from rigorous studies and are relevant to Washington State for use in our model. Therefore, we draw on work by Bohl et al. (2012) for our estimate of the average expected cost of an inpatient hospitalization due to a fall.[80] Bohl and colleagues analyzed Group Health HMO Medicare plans in Washington State to compare the average cost of fallers and non-fallers.[81] We use their estimate, inflated to 2016 dollars, as our average inpatient hospitalization cost, given a hospitalization due to a fall.

We allow parameters to vary in our Monte Carlo analysis as described in Chapter 7, to account for the uncertainty inherent in our estimates. Exhibit 4.3.9 shows the inpatient hospitalization cost and the high and low estimate of the triangle distribution used for our Monte Carlo draws. The high bound of our triangle estimate is drawn from Burns et al. (2016), which draws on national data from the Medical Expenditure Panel Survey (MEPS) and the Medicare Current Beneficiaries Survey (MCBS).[82] This estimate represents the higher end of the cost estimates we found in our literature review. The lower bound of our triangle estimate comes from an analysis from the Washington State Department of Social and Health Services (DSHS)[83] which produced a much lower estimate than what we generally found in the literature.

Exhibit 4.3.9
Inpatient Hospitalization Cost Estimates and Source Literature

Type of estimate	Cost (2016)	Source
Inpatient hospitalization cost	$24,100	Bohl et al. (2012)
Low bound of triangle estimate	$12,442	Washington State DSHS Research and Data Analysis
High bound of triangle estimate	$30,857	Burns et al. (2016)

In addition to inpatient hospitalization costs, the literature indicates that falls incur additional types of health care costs. We calculate the ratio of inpatient cost to other types of health care costs, including emergency department services, outpatient services, and pharmacy/drug costs, and short-term[84] skilled nursing facility placement costs, using the expected costs of these additional health care services from Bohl et al. (2012). These ratios are reported in Exhibit 4.3.10.

Exhibit 4.3.10
Ratios of Other Health Care Costs to Inpatient Hospitalization Cost

Cost type	Ratio
Emergency department	0.211
Outpatient	0.351
Pharmacy/drug	0.072
Short-term skilled nursing facility	0.484

[80] Bohl, A.A., Phelan, E.A., Fishman, P.A., & Harris, J.R. (2012). How are the costs of care for medical falls distributed? The costs of medical falls by component of cost, timing, and injury severity. *The Gerontologist, 52*(5), 664-675.
[81] Ibid.
[82] Burns, E.R., Stevens, J.A., & Lee, R. (2016). The direct costs of fatal and non-fatal falls among older adults—United States. *Journal of safety research, 58*, 99-103.
[83] Estimate provided by the Washington State Department of Social and Health Services (DSHS) Research and Data Analysis Division, November 2017.
[84] We do not estimate the likelihood of long-term nursing facility placement or long-term nursing facility costs due to falls. We are unable to estimate these relationships in Washington State at this time.

We calculate our expected costs of health care due to a fall with the following equations:

$$(4.3.2)\ \Delta FallHCC_y = \frac{\Delta NumberFalls_y \times ChanceHospfromfall_y \times HealthCareFall_y}{(1+dis)^{y-tage}}$$

$$(4.3.3)\ HealthCareFall_y = (HospitalizationCost_y \times (1+HCCEsc)^{y-tage}) \times ratioSkilledNursing \times ratioED \times ratioOutpatient \times ratioPharmacy$$

Finally, we assign health care costs by the payer to participants, taxpayers, and others in society. Due to the fact that older adults (age 65 and over) are eligible for Medicare, the source of health care costs is different for older adults than the rest of the general population. We use the Medicare Current Beneficiaries Survey to calculate the proportion of health care costs by source (Exhibit 4.3.11).

Exhibit 4.3.11

Proportion of Health Care Costs by Source for Individuals Age 65 and Over

	Total cost by perspective			Taxpayer cost by payer		
	Participant	Taxpayer	Other	State	Local	Federal
General health care	16.60%	70.40%	13.00%	4.62%	0%	95.38%
General hospital	0.00%	92.10%	7.90%	0.38%	0%	99.62%
Drug/pharmacy	18.90%	59.10%	22.00%	0.34%	0%	99.66%
Skilled nursing facility	9.20%	83.50%	7.30%	7.01%	0%	92.99%

Note:
WSIPP calculations from the 2013 Medicare Current Beneficiary Survey (MCBS). Centers for Medicare & Medicaid Services (2016). *2013 Medicare current beneficiary survey public use file*.

4.3d Valuing Maternal and Infant Health Outcomes

For maternal and infant health outcomes, we estimate a cost for mothers and for infants where possible. These cost estimates are from a WSIPP analysis of Washington State hospital data linked to singleton births occurring in Washington during the period 2009-2014. For each birth in the dataset, we captured all inpatient hospital costs associated with the mother and with the infant during delivery and over the following year. More information on this analysis can be found in the May 2017 Health Care Technical Appendix.[85]

To model the monetary benefits of changes in maternal and infant health outcomes, we apply the unit change from the standard effect size formula to the costs expected to accrue over a single year. We multiply the average cost of the measured health care resources separately for both the child and mother population (where applicable) by the unit change produced from the program effect size and base rate for that population, adjusted for escalation and discounted to the year of treatment as shown in Equation 4.3.4 below.

$$(4.3.4)\ \Delta HealthCareCosts_{m/i} = \frac{\Delta BirthOutcome_{m/i} \times HealthCareCost_{m/i} \times (1+HCCEsc)^{y-tage}}{(1+dis)^{y-tage}}$$

Exhibits 4.3.12 to 4.3.17 display the average costs and standard errors for mothers and infants separately, during the first year of life, for each birth outcome. These exhibits also display our assumptions about the base rate of the likelihood of each of the outcomes, derived from Washington State data. Exhibit 4.3.18 displays the payer by source information for these costs.

[85] Westley & He (2017).

Exhibit 4.3.12
Preterm Birth Parameters

	General population		Medicaid		Private-pay	
	Mothers	Infants	Mothers	Infants	Mothers	Infants
Average cost for a preterm birth (compared to a non-preterm birth)	$3,078	$24,583	$3,071	$25,267	$3,075	$23,639
Standard error on cost	$77	$551	$123	$873	$101	$705
Year of dollars	2014		2014		2014	
Likelihood of preterm birth^	6.5%		7.5%		5.4%	

Note:
^ Estimates from Washington State Department of Social and Health Services' First Steps Database, (2015). Received April 12, 2017, from the Research and Data Analysis Division, Washington State Department of Social and Health Services.

Exhibit 4.3.13
Low Birthweight (LBW) Birth Parameters

	General population		Medicaid		Private-pay	
	Mothers	Infants	Mothers	Infants	Mothers	Infants
Average cost for LBW birth (compared to a non-LBW birth)	$3,522	$31,299	$3,270	$31,574	$3,714	$31,576
Standard error on cost	$90	$863	$140	$1,435	$120	$1,002
Year of dollars	2014		2014		2014	
Likelihood of LBW birth^	4.9%		5.9%		4.1%	

Note:
^ Estimates from Washington State Department of Social and Health Services' First Steps Database, (2015). Received April 12, 2017, from the Research and Data Analysis Division, Washington State Department of Social and Health Services.

Exhibit 4.3.14
Very Low Birthweight (VLBW) Birth Parameters

	General population		Medicaid		Private-pay	
	Mothers	Infants	Mothers	Infants	Mothers	Infants
Average cost for a VLBW birth (compared to a non-VLBW birth)	$8,592	$145,410	$8,468	$145,379	$8,652	$144,923
Standard error on cost	$372	$4,423	$590	$6,897	$439	$5,282
Year of dollars	2014		2014		2014	
Likelihood of VLBW birth^	0.8%		1.0%		0.6%	

Note:
^ Data for 2013 from Washington State Department of Health, Perinatal Indicators Report for 2014.

Exhibit 4.3.15
Small for Gestational Age (SGA) Birth Parameters

	General population		Medicaid		Private-pay	
	Mothers	Infants	Mothers	Infants	Mothers	Infants
Average cost for an SGA birth (compared to a non-SGA birth)	$234	$3,525	$179	$3,601	$250	$3,079
Standard error on cost	$47	$371	$74	$489	$55	$445
Year of dollars	2014		2014		2014	
Likelihood of SGA birth^	7.1%		7.9%		6.2%	

Note:
^ Estimates from Washington State Department of Social and Health Services' First Steps Database, (2015). Received April 12, 2017, from the Research and Data Analysis Division, Washington State Department of Social and Health Services.

Exhibit 4.3.16
Neonatal Intensive Care Unit (NICU) Parameters

	All infants	Medicaid	Private-pay
Average cost for a NICU admission (compared to no admission to NICU)	$35,132	$40,865	$31,254
Standard error on cost	$721	$1,255	$887
Year of dollars	2014	2014	2014
Likelihood of NICU admission^	7.2%	8.2%	6.3%

Note:
^ Estimates from Washington State Department of Social and Health Services' First Steps Database, (2015). Received April 12, 2017, from the Research and Data Analysis Division, Washington State Department of Social and Health Services.

Exhibit 4.3.17 describes the total costs for a birth by cesarean section, compared to vaginal birth. These estimates are derived from an analysis of MEPS data from 2009 to 2013.

Exhibit 4.3.17
Cesarean Section Parameters

	All mothers	Medicaid	Private-pay
Average cost for a cesarean section (compared to vaginal birth)^	$3,481	$3,021	$3,772
Standard error on cost	$121	$128	$178
Year of dollars	2014	2014	2014
Likelihood of cesarean section#	26.6%	24.0%	28.7%

Notes:
^ WSIPP analysis of pooled annual MEPS data from the 2009-2013 period (five years). Expenditures have been converted to 2014 dollars using medical CPI.
# 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.

Exhibit 4.3.18

Proportion of Maternal and Infant Health Care Costs by Source

	Total cost by perspective			Taxpayer cost by payer		
	Participant	Taxpayer	Other	State	Local	Federal
General	2%	49%	49%	50%	0%	50%
Medicaid	0%	0%	100%	50%	0%	50%
Private-pay	5%	95%	0%	50%	0%	50%

Note:
WSIPP assumptions for participant, taxpayer, and other. Taxpayer cost breakout based on Federal Medical Assistance Percentages for Washington from DHHS ASPE FMAP 2017 Report, Table 1.

WSIPP's benefit-cost model monetizes improvements in health outcomes, in part, with linkages between health conditions and other outcomes to which a monetary value can be estimated. We used Washington State data to estimate the expected effects of individual birth outcomes (preterm, low birthweight, and small for gestational age births) on the likelihood of infant mortality. For each analysis, both the expected effect size and the estimated error 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.

4.4 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.[86] 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.[87] Because the teen birth rate has been trending downward in recent years, we use the most recent data available (2015), which shows a rate of approximately 7.3 teen births per 1,000 women.

4.5 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.[88] 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.

[86] 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.
[87] Retrieved August 2016 from DOH Age-specific Live Birth Rates by Place of Residence. We use the birth numbers for those ages 15-17 from table A10.
[88] 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.[89] 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 in 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.[90] For example, Rosen et al. (2008) 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.[91] 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.[92] 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 information 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 the 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).
- 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.
- 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 along with those lifetime earnings lost because of premature death (mortality).
- Traffic collision costs, to the degree that collisions are estimated to be caused by ATOD (only used in the case of alcohol).

[89] 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.

[90] 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.

[91] 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.

[92] 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.

4.5a 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).[93] 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 Sections describing methods for these topic areas refer back to Section 4.5a.

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 4.5.2 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 4.5b.

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

$$(4.5.1) \quad CPG_y = \left(\sum_{0=1}^{y} O_0 \times P_{(y-0+1)} \right) \times LTP \times S_y \times SF_a$$

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 4.5.2.

$$(4.5.2) \quad 1 = \sum_{y=1}^{100} O_y$$

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 4.5.2. 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 4.1.c. 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:

[93] 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 the recommended maximum weekly or both daily and weekly drinking limits. All outcomes are estimated as dichotomous conditions.

$$(4.5.3) \quad S_y = \frac{LTS_y}{LTS_1}$$

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:

$$(4.5.4) \quad S_y = \frac{LTS_y}{LTS_{tage}}$$

The final term in Equation 4.5.1 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:

$$(4.5.5) \quad SFG_a = \frac{1 - \left(\frac{CondD_a}{(Pop_a \times CP_a)} + \frac{PopD_a - CondD_a}{(Pop_a)} \right)}{\left(1 - \frac{PopD_a}{Pop_a}\right)}$$

In Equation 4.5.5, 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 4.5.1 describes the calculation of the 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:

$$(4.5.6) \quad CPI_y = \frac{\sum_{0=1}^{tage} O_0 \times P_{(y-0+1)}}{\sum_{0=1}^{tage} O_0} \times S_y \times SFI_a$$

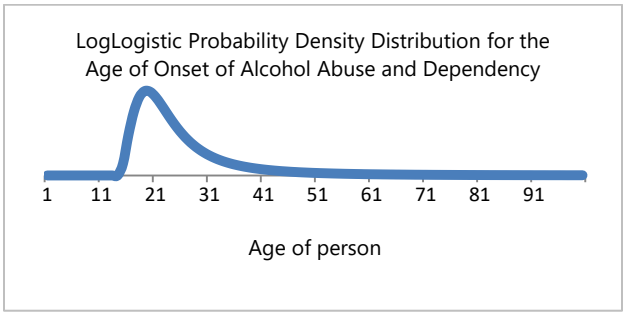
Finally, the survival factors for indicated populations by age group (SFI_a) can be calculated with the following equation:

$$(4.5.7) \quad SFI_a = (SFG_a \times CP_a) + (1 - CP_a)$$

We provide an illustrative example of computing CPG_y in Equation 4.5.1 for disordered alcohol use. Using data from the newer third round of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC III) and definitions from the DSM-V, we applied the methods from Hasin et al., (2007) to compute a probability density distribution for the age of onset of DSM alcohol disorders.[94] We used *@Risk* software to estimate alternative distributions that fit the onset information reported in this nationally representative sample. 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 alcohol use disorder, we computed a log-logistic density distribution; the estimated parameters are reported in Exhibit 4.5.2. Exhibit 4.5.1 plots the estimated distribution, where the sum of annual probabilities equals 1.0.

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

Exhibit 4.5.1



Next, estimates of the persistence of the alcohol disorder, given onset, were computed for alcohol following the methods of Lopez-Quintero, et al.[95] We update the information from the Lopez-Quintero study using NESARC III data. We use the SAS LIFETEST procedure to model the 'survival' of the disorder. Again, we used *@Risk* software to model the best fitting cumulative remission curve and then inverted the result to estimate a persistence curve. A Gamma 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 Gamma parameters are shown in Exhibit 4.5.2 and Exhibit 4.5.3 plots the results.[96]

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

[96] The onset function is shifted with a different parameter in certain instances. When there is a treatment population with a treatment age less than the shift parameter or a program in the general population where there is a measured effect less than the shift parameter, the onset curve is moved to start at the year before the year of treatment or year of measurement as appropriate.

Exhibit 4.5.2

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

	DSM alcohol disorder[1] (a)	Heavy drinking[2] (b)	DSM illicit drug disorder (cannabis)[3] (c)	DSM illicit drug disorder (non cannabis)[4] (d)	DSM illicit drug disorder (opioids)[5] (e)	Regular tobacco smoking[6] (f)
Percentage of population with lifetime DSM disorder, heavy drinking, or regular smoking	29.1%	38.2%	6.3 %	5.6%	2.3%	31.7%
Age of onset: Type of distribution[7]	Log-logistic	Log-logistic	Log-logistic	Log-logistic	Lognormal	Log-logistic
Shift Parameter	14.5238	14.3403	10.5712	13.5864	12.2610	-11.882
Parameter 2	6.5354	6.378	7.5384	7.9854	2.3391	29.31
Parameter 3	2.2368	2.7573	4.188	1.8644	0.8076	16.763
Parameter 4	n/a	n/a	n/a	n/a	n/a	n/a
Remission of DSM disorder, given onset						
Type of distribution[8]	Gamma	Lognormal	Lognormal	Gamma	Gamma	Beta-general
Shift Parameter	0.9522	0.8360	0.3790	0.8680	0.8218	0
Parameter 2	0.4987	1.5571	2.0209	0.6287	0.5840	1.1222
Parameter 3	54.258	2.3444	1.3864	15.668	20.873	2.8754
Parameter 4	n/a	n/a	n/a	n/a	n/a	-2.165
Parameter 5	n/a	n/a	n/a	n/a	n/a	145.55
Percentage of general population consuming substance[9]	68.3%	68.3%	23.2%	11.7%	5.6%	n/a

Notes:

[1] Calculated from NESARC III with lifetime DSM-5 alcohol use disorder.

[2] Calculated from NESARC III. Prevalence is based on the percent exceeding daily/weekly limits in past year. Onset and remission are calculated from the mild classification of DSM-5 alcohol use disorder.

[3] Calculated using NESARC III with lifetime DSM-5 cannabis use disorder.

[4] Calculated as from NESARC III data with lifetime DSM-5 non-cannabis illicit drug substance use disorder. Includes opioids, heroin, sedative, cocaine, stimulant, hallucinogen, inhalant/solvent, club drug, and other drugs.

[5] Calculated as from NESARC III data with lifetime DSM-5 opioid and heroin use disorder.

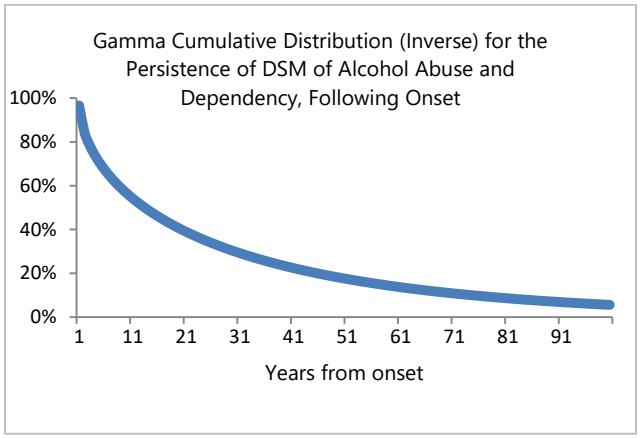
[6] Prevalence is calculated from National Survey on Drug Use and Health (NSDUH) R-DAS online tool, 2-year estimates of Washington State estimates for 2015-2016. Measure is ever daily smoker, variable cduflag. Onset was calculated with NESARC III age at onset of cigarette use. Remission was calculated as persistence of nicotine use disorder among smokers.

[7] Onset curves were calculated using age of onset of a DSM disorder, conditional on having a disorder. We performed an analysis of NESARC-III data, using age of onset for those with disordered conditions. For Log-logistic distributions, Parameter 2 is the scale and Parameter 3 is the shape.

[8] Estimates were constructed following the work of Lopez-Quintero et al. (2011). We used the SAS Lifetest procedure to estimate persistence curves. These values were fitted with @Risk software to estimate distributions; for each disorder, the distribution with the best fit (criterion: lowest root-mean-squared error) was chosen.

[9] Percentage of general population consuming substance estimated from NSDUH R-DAS online tool, 2-year estimates of Washington State estimates for past year use for 2016-2017.

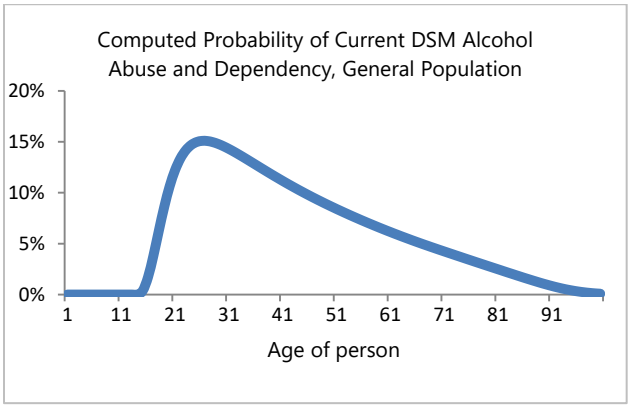
Exhibit 4.5.3



The persistence curve, after multiplying by the survival factor, by year, from the 2016 U.S. 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, 29.1% with sources shown in Exhibit 4.5.2, and after adjusting for survival from the 2016 U.S. 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 4.5.1 and is plotted in Exhibit 4.5.4.

Exhibit 4.5.4



The same procedures just described for alcohol disorders are used for problem 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 4.5.2. As noted, the estimates of the current prevalence of ATOD are central to the benefit-cost model because they become the "base rate" of each ATOD disorder. Program or policy effect sizes are applied to the base rate to determine the change in the number of ATOD "units" caused by the program, over the lifetime following treatment. The general prevalence, shown in Exhibit 4.5.4, is used for programs targeted at the general population, while the persistence curve (after adjustment for survival probabilities and taking into account expected persistence given earlier onset), shown in Exhibit 4.5.3, is used as the base rate for programs that treat people with a current ATOD disorder.

4.5b ATOD Attributable Deaths

WSIPP's model computes mortality-related lost earnings 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 4.5.5 for alcohol, 4.5.6 for tobacco, 4.5.7 for illicit drugs other than cannabis, and 4.5.7 for opioid drugs.

Alcohol. Alcohol-attributable deaths are estimated using a software application called *Alcohol-Related Disease Impact (ARDI)*.[97] ARDI was developed by the U.S. Department of Health and Human Services, Centers for Disease Control (CDC). The application estimates the number of deaths attributable to alcohol causes for each state.

According to the 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 are for the years 2006-10. These are shown in Exhibit 4.5.5.

Exhibit 4.5.5
Alcohol Attributable Deaths by Year, 2006-2010

Age group	Years in age group	Alcohol attributed deaths: Chronic	Alcohol attributed deaths: Acute	Percentage of deaths attributable to DSM alcohol	Percentage 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	1,760,998
20-34	15	12	237	0.50	0.75	1,089	1,369,070
35-49	15	185	260	0.50	0.75	1,338	1,413,666
50-64	15	418	216	0.50	0.75	9,216	1,247,957
65-100	36	344	282	0.50	0.75	35,079	798,384

ARDI estimates deaths related entirely or partially due to particular causes of death. 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. 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.). 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.

[97] Centers for Disease Control and Prevention website.

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. Smoking-attributable deaths are estimated using an on-line software application called *Smoking-Attributable Mortality, Morbidity, and Economic Costs (SAMMEC)*.[98] This data source is also provided through the U.S. Department of Health and Human Services, Center for Disease Control (CDC). SAMMEC estimates the number of deaths attributable to smoking for each state. 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. The latest data available are from 2008.

Exhibit 4.5.6
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	1,217	5,864	768,070
65-74	10	1,582	7,571	413,358
75-84	10	2,262	12,368	251,045
85-100	16	1,456	15,902	111,734

Illicit Drugs and Opioid Drugs. Illicit drug deaths are estimated using Washington State death data from CDC Wonder[99] for the years 2012 to 2016. Opioid drug deaths are estimated using data from the Washington State Department of Health Publication "Opioid-related Deaths in Washington State", 2006–2016 as accessed in April, 2019. We compute average annual drug-attributable deaths in the age groups shown in Exhibit 4.5.7 for other illicit drugs and in Exhibit 4.5.8 for opioids.

Exhibit 4.5.7
Illicit Drug Attributable Deaths by Year, 2012-2016

Age group	Years in age group	Illicit drug attributed deaths	All deaths in state	State population in age group
0-14	14	0	576	1,326,280
15-19	5	15	187	448,523
20-24	5	63	363	477,238
25-34	10	196	895	968,201
35-44	10	212	1,267	911,531
45-54	10	301	3,242	954,459
55-64	10	234	6,836	912,668
65-74	10	56	9,399	583,036
75-84	10	16	12,092	273,760
85-100	16	9	17,618	126,994

[98] Ibid.
[99] Centers for Disease Control and Prevention, National Center for Health Statistics. Multiple Cause of Death Data1999-2016 on CDC WONDER Online Database, released 2018. Data are from the Multiple Cause of Death Files, 1999-2016, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.

Exhibit 4.5.8

Opioid Attributable Deaths by Year, 2012-2016

Age group	Years in age group	Opioid attributed deaths	All deaths in state	State population in age group
0-14	14	0	576	1,326,280
15-24	10	57	550	925,761
25-34	10	143	895	968,201
35-44	10	135	1,267	911,531
45-54	10	179	3,242	954,459
55-64	10	137	6,836	912,668
65-100	36	41	39,110	983,790

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:

$$(4.5.8) \quad AtodD_a = \left((Chronic_a + Acute_a \times AcutePct_a)/Pop_a\right)/Years_a$$

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), $Acute_a$, multiplied by $AcutePct_a$, 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 4.1d.

4.5c Medical Costs, Treatment Costs, and Traffic Accident Damages 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 4.5.9
Health Care Costs for ATOD Disorders

	Alcohol	Cannabis	Opioid drugs	Illicit drugs
Hospital-related parameters				
Average Annual number of disorder FTE hospital events (FY2012-2015)[#]	13,034	4,367	11,450	18,988
Average charge per disorder FTE event (2015 dollars)[^]	$34,698	$17,493	$57,847	$49,129
SD of charge per disorder FTE event	$50,383	$10,871	$101,927	$95,292
Emergency department-related parameters				
Proportion of admissions attributable to substance (2011)	1.06%	0.04%	0.33%	1.01%
Average ED expenses per admission (2015 dollars)	$1,555	$1,555	$1,555	$1,555
SD of average ED expense per admission	$3,587	$3,587	$3,587	$3,587
Treatment parameters				
Annual number treated (2013)	15,046	8,978	11,684	29,868
Average cost per treatment episode (2015 dollars)	$2,156	$2,074	$3,620	$2,783
SD of average cost per treatment episode (2015 dollars)	$2,295	$2,917	$4,617	$3,846

Notes:
[#] FTEHospitalEvent.
[^] 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 for fiscal years 2013-15. CHARS collects information on billed charges of patients, as well as the codes for their diagnoses from hospital inpatient discharge information.[100] We apply the attributable fraction information, described in Section 4.5b, to the CHARS data to estimate the number of hospital events attributable to ATOD.[101] For alcohol related hospital events, we take the average of two potential estimates of the proportion of hospitalizations that could be attributed to alcohol use. Our upper estimate assumes that all events with any code with an Alcohol Attributable Fraction can be attributed to disordered use. Our lower estimate only assumes that only hospital events with any code of AAF of 1 can be attributed to alcohol use.[102]

For the drug use categories, we first followed criteria in Appendix A.1 of the HCUP statistical briefs and the examination of opioid related diagnoses.[103] Guided by these sources, we differed from the Drug Attributable Fraction codes used in Section 4.5b to include the introduction of adverse effect codes, poisoning due to drug use, and maternal use affecting newborns. In instances where the primary code's drug attributable fractionwas less than one, we required a subsequent code to include a code with a drug attributable fraction of one.[104] The illicit drug analysis excluded marijuana codes 304.3 and 305.2, which are the only codes in the marijuana drug use. Opioids are a subset of drug codes focusing on opioids.

The CHARS analysis generates a number of events *FTEHospitalEvents,* as well as the average billed charge per event, *HospCostEvent*, given a stay. These parameters are shown in Exhibit 4.5.10. We also apply a hospital cost-to-charge ratio as described in Section 4.6.

[100] Discharge information is derived from billing systems.
[101] A fully attributable hospital event is one where the Attributable Fraction (AF) equals 1. Hospital events with AFs less than one are summed to create fully attributable hospital events.
[102] For example, the upper bound would include esophageal varices with bleeding (AF < 1), while the lower bound would not. Both would include Alcohol use disorder (AF = 1).
[103] https://www.ncbi.nlm.nih.gov/books/NBK409512/table/sb216.t5/
https://www.hcup-us.ahrq.gov/datainnovations/ICD-10CaseStudyonOpioid-RelatedIPStays042417.pdf
[104] This procedure prevents us from counting instances of, for example, AIDS, when there was no diagnosed sign of drug use.

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:

$$(4.5.9)\quad ExpHospEventsUpperBound = \frac{FTEHospitalEvents}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y-1}^{100} Pop_y}}$$

$$(4.5.10)\quad ExpHospEventsLowerBound = \frac{FTEHospitalEvents}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$(4.5.11)\quad ExpHosp\$ = \frac{ExpHospEventUpperBound + ExpHospEventLowerBound}{2} \times HospCostEvent \times CostRatio$$

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 4.5.10.

Emergency Department Parameters. Emergency department parameters are shown in Exhibit 4.5.10 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 times the total number of emergency room events in Washington. To estimate attribution, we used national data from the HCUP National Emergency Department Sample (NEDS) online tool.[105] Investigations of the number of health events (hospitalization, emergency department visits, death) rely on the ICD-9 clinical classification system. WSIPP reviewed literature on ICD-9 coding practices and assignment of attribution. We calculated the proportion of admissions attributable to substances as the percent of all ED visits in the NEDS 2014 sample for which an eligible ICD-9 code or E-code associated with the admission was the primary diagnosis of the admission. Codes for alcohol were taken from White et al. (2018), opioids from Weiss et al (2017), illicit drugs from Sevigney & Caces (2018) (excluding codes for marijuana), and marijuana from Hall et al (2018).

[105] HCUPnet, Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. https://hcupnet.ahrq.gov/. For more information about HCUP data see http://www.hcup-us.ahrq.gov/

The total number of emergency department visits in Washington during 2017 is entered in Exhibit 4.5.10. These data come from the Washington State Hospital Association.[106] We then apply the proportion of admissions attributable to substances just described; for example, for DSM alcohol disorders, we apply the 1.06% factor calculated from NEDS to the number of visits in Washington to determine the number of alcohol-related emergency room visits in Washington. As with hospital events, we compute the upper and lower bounds by dividing by the current annual prevalence of DSM disorders in the general population (upper bound) or the current level of any alcohol use not just DSM disorders in the general population (lower bound). We then apply a cost per emergency department event, *EDCostEvent,* and an emergency department cost-to-charge ratio. The average and standard error of the cost per emergency department visit is taken from the Medical Expenditure Panel Survey (MEPS) of the U.S. Department of Health & Human Services.[107] 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:

$$(4.5.12)\quad ExpEDEventsUpperBound = \frac{TotalEDVists \times CausationFraction}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y-1}^{100} Pop_y}}$$

$$(4.5.13)\quad ExpEDEventsLowerBound = \frac{TotalEDVists \times CausationFraction}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$(4.5.14)\quad ExpED\$ = \frac{ExpEDEventsUpperBound + ExpEDEventsLowerBound}{2} \times EDCostEvent \times CostRatio$$

Treatment Parameters. For the costs of admissions to treatment, WSIPP was supplied with numbers by the Washington Department of Social and Health Services (DSHS). The number of admissions comes from the Treatment and Assessment Report Generation Tool (TARGET) database for FY 2013.[108] 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:

$$(4.5.15)\quad ExpTreatmentEvents = \frac{TotalTreatmentEvents}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y-1}^{100} Pop_y}}$$

$$(4.5.16)\quad ExpTreatment\$ = ExpTreatmentEvents \times TreatmentCostEvent$$

[106] Number is from the American Hospital Association survey of community hospitals as provided by Matt Shevrin of the Washington State Hospital Association in personal correspondence, September 14, 2019.
[107] Analysis of 2015 MEPS data. Average annual ED Cost of those with a visit. For more on MEPS, see Section 4.3.
[108] Information from the TARGET database was provided via personal communication with Kevin Campbell, DSHS, May 12, 2016. Data changes in the Washington State behavioral health system have led to a current gap in the data. We look forward to updating this number as new information becomes available.

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).[109] To estimate the proportion of all crashes that are reported by police out of total crashes, we use national estimates produced by Blincoe et al., (2002).[110] Data from Blincoe provide an estimate that 56.7% of all crashes are reported by police.[111] 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 4.5.11).

Exhibit 4.5.10
Calculation of Average Property Costs from Alcohol-Caused Traffic Collisions

Collision category	Unit price in 2000 dollars	Total alcohol caused incidence	Percentage 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

Note:
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:

$$(4.5.17)\quad ExpTrafficCollisionsUpperBound = \frac{TotalTrafficCollisions \times CausationFraction}{\frac{\sum_{y=1}^{100} CP_y \times Pop_y}{\sum_{y-1}^{100} Pop_y}}$$

$$(4.5.18)\quad ExpTrafficCollisionsLowerBound = \frac{TotalTrafficCollisions \times CausationFraction}{CurrentUse\% \times \sum_{y=1}^{100} Pop_y}$$

$$(4.5.19)\quad ExpTrafficCollison\$ = \frac{ExpTrafficCollisionsUpperBound + ExpTrafficCollisionsLowerBound}{2} \times TrafficCostEvent$$

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 the 2008-2011 Medical Expenditure Panel Survey. As explained in more detail in Section 4.6, MEPS data include a representative sample of NHIS households with additional detail collected on individual health care utilization and medical expenditures. We follow the methodology outlined by Xu, et al., (2015)[112] in constructing a two-part model that examines smoking-attributable health care spending controlling for sociodemographic characteristics and other health-related behaviors and attitudes.

[109] Washington State Department of Transportation. (n.d.). *2011 Washington State collision data summary*. Olympia, WA: Author, Table 8.
[110] 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.
[111] Ibid, table 3.
[112] 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.

Two separate models were included in this analysis—a *prevention* model that estimated costs for non-smokers[113] 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-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 the 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.[114]

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 health care expenditures are estimated conditioned on entering the specified smoking status. The dependent variable, total health care 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 in Appendix III in Exhibits III.1 and III.2.

After deriving adjusted values for the overall effect of smoking on health care 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 4.5.12 shows the average annual cost and incremental cost by year for prevention and treatment populations.

Exhibit 4.5.11

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
Age-based cost of disorder for each year from measurement age	$21.68	$7.84
Standard error on additional cost	$1.64	$3.15

4.5d Human Capital Outcomes Affecting Labor Market Earnings via ATOD-Caused Morbidity

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 Section 4.1d, WSIPP's model uses national earnings data from the U.S. Census Bureau's Current Population Survey. The CPS data used in this analysis represent average earnings of all people, both workers and non-workers at each age.

For each person at each age, total CPS earnings can be viewed as a weighted sum of people who have never had an ATOD disorder, plus those that are currently disordered, plus those that were formerly disordered, 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:

[113] Note: non-smokers are defined as individuals that smoked less than 100 cigarettes during a lifetime.
[114] United States. (2012). *Preventing tobacco use among youth and young adults: A report of the Surgeon General*. Rockville, MD: U.S. Dept. of Health and Human Services, Public Health Service, Office of the Surgeon General.

$$(4.5.20)\quad EarnC_y = \frac{ModEarnPop}{\left((EarnGN)\times\left(1-\left(CP_y+\left(\sum_{o=1}^{y}(O_o\times LTP)-CP_y\right)\right)\right)+(EarnGF)\times\left(\sum_{o=1}^{y}(O_o\times LTP)-CP_y\right)+CP_y\right)}$$

The numerator in the above equation contains the selected earnings population as described in Section 4.1. uses our modified CPS earnings as described in Section 4.1 and shown in Equation 4.2.2. This will typically be *ModEarnAll*, or the average compensation of the population in Washington

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

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 the employment and earnings of workers. We use CPS earnings over the last business cycle 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 Section 4.2d. 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 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 4.5.9. 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. For clarity, Exhibit 4.5.9. also presents the expected value of the ratio for each distribution.

Exhibit 4.5.12
Labor Market Earnings Parameters for ATOD Disorders

		DSM alcohol disorder	Problem alcohol use	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	Expected Ratio (non-disordered to disordered)	1.255	1.116	1.136	1.136	1.136	1.118
	Distribution type	Gamma	Lognormal	Gamma	Gamma	Gamma	Normal
	Alpha/mean	48.687	-0.4744	47.337	47.337	47.337	1.08671
	Beta/standard deviation	0.01059	0.13262	0.00510	0.00510	0.00510	0.02814
	Shift	0.74332	0.49361	0.89631	0.89631	0.89631	NA
Gain in labor market earnings for former users vs current disordered users, probability density distribution parameters	Expected Ratio (non-disordered to disordered)	1.255	1.116	1.136	1.136	1.136	1.118
	Distribution type	Gamma	Lognormal	Gamma	Gamma	Gamma	Normal
	Alpha/mean	48.687	-0.4744	47.337	47.337	47.337	1.08671
	Beta/standard deviation	0.01059	0.13262	0.00510	0.00510	0.00510	0.02814
	Shift	0.74332	0.49361	0.89631	0.89631	0.89631	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:

$$(4.5.21)\ PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta ATOD_y \times \left(1 - \sum_{o=1}^{y} O_o\right) \times EarnGN \times EarnC_y\right) + \left(\Delta ATOD_y \times \left(1 - \left(1 - \sum_{o=1}^{y} O_o\right)\right) \times EarnGF \times EarnC_y\right)}{(1 + dis)^{(y-tage+1)}}$$

Where *ΔATODy* 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:

$$(4.5.22)\ PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta ATOD_y \times EarnGF \times EarnC_y\right)}{(1 + dis)^{(y-tage+1)}}$$

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 the probability that a person dies as the result of the disorder given that they have the disorder at that particular age and includes the value of the mortality risk reduction due to ATOD. For more on the VSL calculations, see Section 4.1.d.

4.5e 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 the expected effect size and the estimated error 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.

4.5f 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 update our earlier original NESARC analysis using the latest round of NESARC data. Our estimates and sources for these early initiation parameters are described in Exhibit 4.5.15. These estimates are treated as links between measured early initiation and later disordered use. We apply our standard links procedures as described in Section 3.4.

Exhibit 4.5.13
Early Initiation Parameters

	Alcohol (a)	Cannabis (b)	Illicit drugs (non cannabis) (c)	Regular tobacco smoking (d)
Prevalence of substance use: [1]				
By end of middle school	23.5%	13.9%	9.8%	9.1%
By end of high school	58.5%	43.6%	18.9%	23.8%
D-cox effect size (ES) between early initiation and later disorder[2]				
By end of middle school	0.582	0.987	1.184	0.676[3]
By end of high school	0.759	1.748	1.627	1.181[3]
Standard error on D-cox ES between early initiation and later disorder[2]				
By end of middle school	0.032	0.049	0.064	0.007[3]
By end of high school	0.024	0.051	0.046	0.007[3]

Notes:
[1] Miech, R. A., Schulenberg, J. E., Johnston, L. D., Bachman, J. G., O'Malley, P. M., & Patrick, M. E. (December 17, 2018). "National Adolescent Drug Trends in 2018." Monitoring the Future: Ann Arbor, MI. Retrieved October, 2019 from http://www.monitoringthefuture.org. 8th grade and 12th grade 2018 numbers from Table 1 reported.
[2] Analysis of NESARC III data. We looked at the odds ratio of the likelihood of later disordered use for those who began using a substance (Alcohol, Marijuana, Other Illicit Drugs, Opioids) in either middle school or high school as compared to those who did not initiate early (including those who never initiate). This analysis controlled for respondent demographics of age, gender, and race/ethnicity, feelings of parent connection and trauma, as well as parent/adult in home behaviors including parent substance and mental health. 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.
[3] Analysis of NSDUH data from 2002-2016. We looked at the odds of ever being an everyday smoker for someone who initiated smoking in either middle school or high school as compared to those who did not initiate early (including those who never initiate). This regression controlled for year, age of respondent, sex, and race. Although the NESARC III analysis provides a larger variety of early life controls, respondents to the survey must have smoked 100 cigarettes over the course of their lifetime to be asked about their early initiation. 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.

4.6 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 attention-deficit/hyperactivity disorder (ADHD),

depression, anxiety, disruptive behavior, internalizing behavior, externalizing behavior, and post-traumatic stress disorder (PTSD). The category of disruptive behavior covers the DSM categories of oppositional defiant disorder and conduct disorder. Obviously, there are other recognized mental health disorders. It is anticipated that the 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 Section 4.5 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.[115] For example, Insel (2008) summarizes findings indicating the total cost of serious mental illness in the U.S. 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% due to disability payments.[116] 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 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).
- ✓ 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, along with the lifetime earnings which are lost because of this premature death (mortality).
- ✓ 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.

4.6a Mental Health Parameters.

WSIPP's mental health model is driven by 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.

4.6b 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 alcohol, tobacco, or other drugs (ATOD) disorders; please see Section 4.5b for formulas and detailed descriptions.

---

[115] 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; 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; and 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.

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

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 4.6.1 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 4.6c in this Chapter and displayed in Exhibit. 4.6.2.

Exhibit 4.6.1

Input Parameters for the Epidemiology of Mental Health Disorders[1]

	ADHD	Depression	Anxiety	Internalizing Behaviors	Disruptive Behaviors	Externalizing Behaviors	DSM PTSD
Percent of population with lifetime DSM disorder[2]	8.0%	23.0%	31.9%	6.1%	14.9%	23.1%	8.7%
Age of onset							
Type of distribution[3]	Laplace	Log-normal	Log-normal	Beta	Beta	Log-normal	Log-logistic
Parameter 1 (Shift)	0	0	0	0	0	0	0
Parameter 2	7.099	3.5755	2.2282	2.9464	2.72010	2.33110	23.815
Parameter 3	1.681	0.7035	0.6069	1.05570	1.41840	0.49019	2.2680
Parameter 4				0	0		
Parameter 5				18	18.028		
Remission of DSM disorder, given onset							
Type of distribution[4]	Log-normal	Beta	Beta	Beta	Log-logistic	Log-normal	Beta
Parameter 1 (Shift)	0	0	0	0	0.41682	-0.26750	0
Parameter 2	3.2391	0.5077	0.83011	0.56643	6.03870	2.78410	0.72016
Parameter 3	1.50970	2.4017	2.00780	2.82730	1.45870	1.42440	1.38730
Parameter 4		0.9994	0	0			-1.66910
Parameter 5		128.35	196.73	166.33			180.78

Notes:

[1] We follow the methodology used to analyze the NCS-R in Kessler, R.C., Berglund, P., Delmer, O., Jin, R., Merikangas, K.R., & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 593-602. We produced our estimates using 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 the prevalence of mental illnesses in the U.S. population.
We differ from Kessler in several places. The estimate for disruptive behavior is an average of the reported risk for oppositional-defiant disorder and conduct disorder. Internalizing and externalizing were constructed using non-hierarchical factor diagnoses described in Kreuger, R (1999). The Structure of Common Mental Disorders. *Archives of General Psychiatry,* 56(10), 921-926. Internalizing consists of major depressive episode, dysthymia, and generalized anxiety disorder. Externalizing consists of conduct disorder, oppositional-defiant, intermittent explosive, and ADHD.

[2] These numbers represent the percent of the population who will develop the disorder in their lifetime, calculated from the lifetime onset tables described above at 75 for Depression, Anxiety, ADHD, Disruptive Behavior, and PTSD. For internalizing and externalizing, the lifetime prevalence was measured at age 18.

[3] Again we follow the methodology used in Kessler et al. (2005). All age of onset distributions were fit with life tables created using the methods that generated Table 3 in the paper. We estimated probability density distributions for the age of onset of each of the 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. For disruptive behavior, we combined the onset curves from oppositional defiant disorder and conduct disorder. Parameters are listed in the order in which they are entered into Excel formulas (with the shift parameter as an addition before the formula).

[4] We identified persons with a lifetime diagnosis of the relevant disorder in the NCS-R. 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 (as measured at the time of the survey). 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.

4.6c Mental Health Attributable Deaths

WSIPP's model computes mortality-related lost earnings 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 4.6.2 below. For both of these disorders, we assume that a proportion of deaths by suicide are caused by mental illness.

Exhibit 4.6.2

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	Percent of suicides attributable to depression	Percent 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 U.S. 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*.[117] 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 U.S. 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 4.6.2, 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.

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:

$$(4.6.1)\quad MHD_a = \big((Suicides_a \times MHSuicidePct)/Pop_a\big)/Years_a$$

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, $Suicide_a$, by 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 4.1d.

[117] Centers for Disease Control and Prevention website.

4.6d 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 the expected effect size and the estimated error 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.

4.6e 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 Section 4.2, WSIPP's model uses national earnings data from the U.S. Census Bureau's Current Population Survey (CPS). The CPS data used in this analysis represent the average earnings of all people, both workers and non-workers at each age.

Exhibit 4.6.3

Labor Market Earnings Parameters for Mental Health Morbidity and Mortality

		Depression	Anxiety	PTSD
Gain in labor market earnings for never used vs. current disordered users, probability density distribution parameters	Expected ratio (mental health condition vs. no condition)	1.213	1.258	1.200
	Distribution type	Gamma	Gamma	LogNormal
	Alpha/mean	59.063	46.851	-0.92055
	Beta/standard deviation	0.00676	0.01072	0.1669
	Shift	0.79839	0.73366	0.77784
Gain in labor market earnings for former users vs. current disordered users, probability density distribution parameters	Expected ratio (mental health condition vs. no condition)	1.213	1.258	1.200
	Distribution type	Gamma	Gamma	LogNormal
	Alpha/mean	59.063	46.851	-0.92055
	Beta/standard deviation	0.00676	0.01072	0.1669
	Shift	0.79839	0.73366	0.77784

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 the 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:

$$(4.6.2)\quad EarnC_y = \frac{ModEarnPop}{\left((1+EarnGN)\times\left(1-\left(CP_y+\left(\sum_{o=1}^{y}(O_o\times LTP)-CP_y\right)\right)\right)+(1+EarnGF)\times\left(\sum_{o=1}^{y}(O_o\times LTP)-CP_y\right)+CP_y\right)}$$

The numerator in the above equation contains the selected earnings population as described in Section 4.1. uses our modified CPS earnings as described in Section 4.1 and shown in Equation 4.2.2. This will typically be *ModEarnAll*, or the average compensation of the population in Washington.

The denominator in Equation 4.6.2 uses the epidemiological variables described above: age of onset probabilities, *Oo,* 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 Section 2.1 of Chapter 2.

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 the employment and earnings of workers. We use CPS earnings over the last business cycle 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 Section 4.2d. 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 4.6.3. 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:

$$(4.6.3)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta MH_y \times \left(1 - \sum_{o=1}^{y} O_o\right) \times EarnGN \times EarnC_y\right) + \left(\Delta MH_y \times \left(1 - \left(1 - \sum_{o=1}^{y} O_o\right)\right) \times EarnGF \times EarnC_y\right)}{(1 + Dis)^{(y-tage+1)}}$$

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:

$$(4.6.4)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta MH_y \times EarnGF \times EarnC_y\right)}{(1 + Dis)^{(y-tage+1)}}$$

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 the 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 the functioning of those individuals, not necessarily to relieve their mental illness itself. Whereas for the mental health conditions of depression, anxiety, and PTSD, we estimate changes in labor market earnings via the impact of the program on the mental health condition (as described above), in evaluations of intervention programs for those with serious mental illness, the best measure of labor market participation is often employment rather than serious mental illness itself. Therefore, we estimate changes in labor market earnings for individuals with serious mental illness only in cases where employment is measured. We apply the calculated unit change in employment resulting from the program to the expected earnings for a population with serious mental illness, *EarnSMI*. This factor is described in Section 4.2c.

4.6f Medical Costs from Mental Health

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 4.6.4. They were computed from an analysis of data from the federal Medical Expenditure Panel Survey (MEPS).

Exhibit 4.6.4

Annual Expected Costs of Mental Health Conditions

		DSM ADHD	DSM depression	DSM anxiety	Internalizing	Disruptive behavior	Externalizing	DSM PTSD
Child (age 1-17)	Annual $	$1,084	$938	$938	$657	$1,817	$1,122	$1,817
	SD	$316	$566	$566	$346	$622	$419	$622
	Year of $	2015	2005	2005	2005	2005	2005	2005
Adult	Annual $	$1,084	$1,763	$553	$657	$1,817	$1,122	$1,817
	SD	$316	$915	$526	$346	$622	$419	$622
	Year of $	2015	2011	2011	2005	2005	2005	2005

Estimates for Mental Health Disorders. MEPS is a nationally representative large-scale survey of American families, medical providers, and employers who report on health care service utilization and associated medical conditions, costs, and payments. Additional information about MEPS can be found in Section 4.3**.**

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-reported depression and anxiety[118] and linked these individuals to health care expenditure information from the 2008-2009 MEPS survey.[119] Post-traumatic stress disorder (PTSD) is not identified for respondents in the NHIS or MEPS. In order to estimate costs for patients with PTSD, we used a finding by Ivanova et al. (2011) that the incremental costs of PTSD are 8% higher than that for major depressive disorder.[120]

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 depression and anxiety in children, and the "conduct problems" scale to estimate costs for disruptive behavior. We also estimate the costs associated with two aggregate scales. "Internalizing" problems are identified using the sum of the emotional and peer scales, and "externalizing" problems

[118] During the past 12 months, have you been frequently depressed? During the past 12 months, have you been frequently anxious?
[119] Center for Disease Control and Prevention. *National Health Interview Survey*.
[120] Ivanova, J., Birnbaum, H., Chen, L., Duhig, A., Dayoub, B., Kantor, E., . . . Phillips, G. (2011) Cost of post-traumatic stress disorder vs major depressive disorder among patients covered by Medicaid or private insurance. *American Journal of Managed Care*, 17(8), e314-e323.

are identified by using the sum of the conduct and hyperactivity scales. Responses for children in the sample child core questionnaire are linked to subsequent health care expenditures in the 2004-2005 MEPS survey.

Recent MEPS survey rounds identify ADHD among children aged five through 17 as a "priority condition." Survey respondents are asked if each child had ever been diagnosed with ADHD. We were, therefore, able to use more recent 2015 MEPS survey data to estimate the medical costs associated with ADHD.

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.[121] 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 health care costs (e.g., age, presence of chronic illnesses, health insurance status, education).

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.[122] 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,[123] 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. Our regression results can be found in Appendix III at the end of this document.

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 the 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:

$$(4.6.5)\quad PV\Delta HC = \sum_{y=tage}^{100} \frac{\left(\Delta HCOutcome_y \times HCCostSMI\right)}{(1+Dis)^{(y-tage+1)}}$$

In Equation 4.6.5, *HCCostSMI* is estimated from the sources listed in Exhibit 4.6.6. 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 4.6.5.

[121] Glick, H. (2007). *Economic evaluation in clinical trials*. Oxford: Oxford University Press.
[122] Buntin, M.B., & Zaslavsky, A.M. (2004). Too much ado about two-part models and transformation? *Journal of Health Economics,* 23(3), 525-542.
[123] Glick (2007) and Buntin & Zaslavsky (2004).

Exhibit 4.6.5

Expected Costs of Health Care Resources Used by Individuals with Serious Mental Illness

	Emergency department	Hospital (general)	Hospital (psychiatric)
Annual $	$1,848	$15,145	$21,356
SD	$2,920	$19,283	$19,709
Year of $	2015	2015	2012
Annual percent of seriously mentally ill adults using resource	42.2%	24.3%	8.3%

Exhibit 4.6.6

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 2015 MEPS data; sample-weighted average cost of ED visits for those classified as SMI (Schizophrenia and other psychotic disorders), conditional on having at least one ED visit in the year.	WSIPP analysis of 2015 MEPS data; of those classified as having SMI, proportion who were treated in the emergency room at least once in the past year.
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 2015 MEPS data; of those classified as having SMI, proportion who were admitted to the hospital.
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 the 2012 number of clients residing in the state mental hospitals, provided by DSHS Research and Data Analysis division, divided by the estimated total population of seriously mentally ill individuals in Washington.

4.7 Valuation of Health Conditions — Obesity and Diabetes

WSIPP models health conditions (currently limited to diabetes and obesity) following the same general analytic procedures described in Section 4.8 for mental health disorders. Readers can refer to that section for additional detail. WSIPP's model uses an incidence-based costing approach to look at the long-term economic implications of diabetes and obesity, as described in Section 4.7d.

Finally, we also model the value of other related outcomes when health conditions (such as diabetes and obesity) are not directly measured by outcome evaluations. For example, we examine the economic implications of weight loss through its causal link to diabetes. These relationships are discussed in Section 4.3f.

4.7a Health Condition 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.

The methods used to compute the current prevalence of health conditions are the same as those used to compute the current prevalence of alcohol, tobacco, and other drugs (ATOD) disorders; please see Section 4.5b 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 4.7.1 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 later in this section.

In Exhibit 4.7.2, 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 4.7.1. Estimates for obesity were obtained using the NLSY—National Longitudinal Survey of Youth.[124] 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.[125] 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, the 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.[126] 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.

[124] Bureau of Labor Statistics. *National Longitudinal Surveys*.
[125] National Longitudinal Surveys. *National Longitudinal Survey of Youth 1979*.
[126] Centers for Disease Control and Prevention. *CDC Growth Charts*.

Exhibit 4.7.1
Input Parameters for the Epidemiology of Health Conditions

	Type 2 diabetes	Obesity
Percentage of population with condition at any point in lifetime	37%[1]	58.4%[2]
Percentage 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:

[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 U.S. *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 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.

4.7b Deaths Attributable to Health Conditions

WSIPP's health conditions model computes mortality-related lost earnings 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. (2002)[127] 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.[128] 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 on the mortality rate in 317,875 men and women over 20 years. We computed a weighted average of the results from these two 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 4.1d.

$$(4.7.1) \quad DiabDR_y = \frac{\frac{DiabD_a}{Pop_a}}{Years_a}$$

4.7c 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 Section 4.2, WSIPP's model uses national earnings data from the U.S. Census Bureau's Current Population Survey (CPS). The CPS data used in this analysis represent the average earnings of all people, both workers and non-workers at each age.

[127] 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.

[128] 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 U.S. adults. *New England Journal of Medicine, 341*(15), 1097-1105 and Calle, E.E., Teras, L.R., & Thun, M.J. (2005). Obesity and mortality. *New England Journal of Medicine, 353*(20), 2197-2199.

Exhibit 4.7.2
Labor Market Earnings Parameters for Health Morbidity

		Type 2 diabetes: Age 50+	Obesity
Gain in labor market earnings for never used vs. current disordered users, probability density distribution parameters	Distribution type	Normal	Normal
	Mean	1.19125	1.06643
	Standard deviation	0.05387	0.04109
Gain in labor market earnings for former users vs. current disordered users, probability density distribution parameters	Distribution type	Normal	Normal
	Mean	1.19125	1.06643
	Standard deviation	0.05387	0.04109

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:

$$(4.7.2)\quad EarnC_y = \frac{ModEarnPop}{\left((1+EarnGN)\times\left(1-\left(CP_y+\left(\sum_{o=1}^{y}(O_o\times LTP)-CP_y\right)\right)\right)+(1+EarnGF)\times\left(\sum_{o=1}^{y}(O_o\times LTP)-CP_y\right)+CP_y\right)}$$

The numerator in the above equation contains the selected earnings population as described in Section 4.1. uses our modified CPS earnings as described in Section 4.1 and shown in Equation 4.2.2. This will typically be *ModEarnAll*, or the average compensation of the population in Washington

The denominator in Equation 4.7.6 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 Appendix II.

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 Section 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 the employment and earnings of workers. We use CPS earnings over the last business cycle 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 Section 4.2d. 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 4.7.2. 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:

$$(4.7.3)\quad PV\Delta Earn = \sum_{y=tage}^{65} \frac{\left(\Delta H_y \times \left(1 - \sum_{o=1}^{y} O_o\right) \times EarnGN \times EarnC_y\right) + \left(\Delta H_y \times \left(1 - \left(1 - \sum_{o=1}^{y} O_o\right)\right) \times EarnGF \times EarnC_y\right)}{(1 + Dis)^{(y-tage+1)}}$$

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:

$$(4.7.4)\quad PV\Delta Earn = \sum_{y=Tage}^{65} \frac{\left(\Delta H_y \times EarnGF \times EarnC_y\right)}{(1 + Dis)^{(y-tage+1)}}$$

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 the probability that a person dies as the result of the disorder given that they have the disorder at that particular age.

4.7d Medical Costs for Specific Health Conditions.

Exhibit 4.7.3 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 4.7.3
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	$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 with pain treatment may be related to multiple underlying conditions.

To estimate the total health care 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 in Section 4.6f. 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, the cost of illness models are based on public data available in the federal Medical Expenditure Panel Survey (MEPS). Additional information about MEPS is provided at the beginning of Section 4.7. 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 U.S. In 2012, over 22.3 million Americans were diagnosed with diabetes (7% of the U.S. 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.[129] 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.[130] 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 A.III.10 shows the results for our two-part model of health care 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.

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.[131] 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, the 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.

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 U.S. 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

[129] American Diabetes Association. (2013). Economic costs of diabetes in the U.S. in 2012. *Diabetes Care,* 36(4), 1033-46.
[130] Medical Expenditure Panel Survey (MEPS).
[131] An, R. (2015). Health care expenses in relation to obesity and smoking among U.S. 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 U.S. 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 U.S. 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; and 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.

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. [132]

Exhibit 4.7.4 displays the assumptions and estimated annual costs we use when computing nursing home costs.

Exhibit 4.7.4
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.[133] According to this survey, the median intermediate cost for a semi-private room was $253 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 Exhibits A.I.1 to A.I.3 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 the 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).[134] 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%).[135] These payer by source numbers are presented in Exhibit 4.7.5.

Exhibit 4.7.5
Proportion of Obesity and Diabetes Health Care Costs by Source

	Total cost by perspective			Taxpayer cost by payer		
	Participant	Taxpayer	Other	State	Local	Federal
Obesity: Under age 65^	12.77%	28.24%	58.98%	28.20%	0.00%	71.80%
Obesity: Age 65 and over^	12.67%	70.02%	17.31%	2.49%	0.00%	97.51%
Diabetes: Under age 65^	11.53%	39.21%	49.26%	21.88%	0.00%	78.12%
Diabetes: Age 65 and over^	11.37%	73.02%	15.61%	3.53%	0.00%	96.47%
Nursing home#	33.71%	62.38%	3.91%	41.42%	0.00%	58.58%

Notes:
^ WSIPP calculation from 2013 Medical Expenditure Panel Survey data.
# Cost by perspective calculated from the National Nursing Home Survey 2004.

[132] Preston, S.H., Fishman, E., & Stokes, A., (2014). *Lifetime probability of developing diabetes in the United States.* PSC Working Paper Series, WPS 14-4.
[133] 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.
[134] ICPSR04651-v1. Hyattsville, MD: U.S. 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.
[135] Centers for Disease Control and Prevention. *Note on coding payment sources in the NNHS 2004.*

4.7e Linkages: Health Outcomes to Other Outcomes

WSIPP's benefit-cost model monetizes improvements in health outcomes, in part, with linkages between health conditions 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. For each analysis, both the expected effect size and the estimated error 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.

4.8 Valuation of K–12 Education Outcomes

In valuing most K–12 education outcomes (i.e., standardized test scores, high school graduation), we use a human capital approach, as described in Section 4.2. This section describes the inputs (Section 4.8a) and computational procedures (the subsequent sections) 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.

4.8a Education Parameters

Evaluations of education and other programs or policies often assess outcome measures such as student test scores, 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 4.8.1. This section lists the individual inputs and their data sources.

Exhibit 4.8.1
General K–12 Education Parameters

		All students	Low-income students
State high school graduation rate		0.781	0.680
Cost of a year of education (2017 dollars) for a student in regular education		$9,585	$11,299
Cost of a year of education (2017 dollars) for a student in special education		$20,571	$22,285
Percentage of students using special education		0.141	0.198
Average numbers of years in special education, for those who receive it		9.86	10.20
Average age of first entry into special education		6.20	6.50
Percentage of students retained for at least one year		0.108	0.119
Average number of years retained, for those retained		1	1
Multiplier for human capital economic externalities of education	Max	0.42	0.42
	Mode	0.37	0.37
	Min	0.125	0.125
Gain in earnings for a 1SD increase in test scores	Mean	0.0978	0.0978
	SE	0.0313	0.0313
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).[136] 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 low-income students.[137] In addition, WSIPP uses a lower predicted high school graduation rate for the juvenile offender population.[138] 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.

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.[139] The cost of K–12 education for a low-income student is calculated by adding the per low-income student amount calculated from the compensatory education expenditures category.

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.[140] 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 2017–18 (14.1%).[141] WSIPP calculates the rate at which low-income students receive special education using information provided by Washington State.[142] 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 first used.[143]

[136] Office of Superintendent of Public Instruction. (2015). *Graduation and Dropout Statistics Annual Report: Appendix A*. Olympia, WA.
[137] 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.
[138] 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. 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. These numbers are in line with numbers calculated from Table 4 of the December 2016 Juvenile Justice Standardized Report *Education and Workforce Outcomes of Juvenile Justice Participants in Washington State* authored by the Education Research & Data Center at the Office of Financial Management.
[139] The cost of regular education estimate is from Office of Superintendent of Public Instruction. (2017). 2016-2017 *Financial reporting summary: Washington State School Districts, Charter, Tribal Schools, and Educational Service Districts*. Olympia, WA: Dawn-Fisher, Lisa, Table.
[140] 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).
[141] Office of Superintendent of Public Instruction. *Washington State Report Card*.
[142] Information from S. Grummick, OSPI (personal communication, September 6, 2018).
[143] The average number of years of special education is from S. Grummick, OSPI (personal communication, September 6, 2018). The average age of first entry in special education is developed from information from L. Diao (personal communication, September 26, 2018).

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. Grade retention estimates are based on data provided by Washington's Office of Superintendent of Public Instruction (OSPI). OSPI provided information on the retention rates for students enrolled in all grade levels from 2011-2019. Our low income student estimate is based on the retention data provided for receiving free- or reduced-price lunch students.

The estimate for the "average number of years retained, for those retained" reflects the average number of years retained for all students, rounded to the closest year.[144] We estimate that students will only be retained for a single year. Since their data did not follow any cohort through the duration of their academic career, the estimate for the "percent of student retained at least one year" was calculated from the sum of the average probability of being held back in each grade.

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 4.8.1. There is a fairly large economic literature on this topic, summarized in a chapter by McMahon in Brewer (2010).[145] Analysts have studied the degree to which growth in the private returns to human capital produces spillover economic gains to the rest of an economy. The low value we use is the estimate contained in Acemoglu & Angrist (2000).[146] The modal value is the estimate used in Belfield, Hollands, and Levin (2011).[147] The high parameter is contained in Bretton (2010).[148] 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 by 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 4.8.2.

Exhibit 4.8.2
Proportion of Marginal Education Costs by Source

	State	Local	Federal
Regular education^	0.7166	0.2115	0.0719
Special education#	0.8668	0.000	0.1332

Notes:
^ Washington State Office of the Superintendent of Public Instruction, *2016-2017 Financial Reporting Summary*, Table 3.
# 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.

[144] Less than 2 percent of all students enrolled in kindergarten through 9th grade are held back from more than a year. This number increases slightly in the tenth grade and highest for twelfth graders. Based on consecutive retention, it appears that this is because students in high school are held back in the same grade repeatedly, rather than because prior retention predicts subsequent retention. This strengthen our belief that changed in student retention are likely to only impact the grade the observed grade and not affect later student retention.
[145] McMahon, M. (2010). The external benefits of education. In D.J. Brewer, & P.J. McEwan (Eds.) *Economics of education*. Oxford, UK: Academic Press.
[146] Acemoglu, D., & Angrist, J. (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual,* 15, 9-59.
[147] 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.
[148] Breton, T.R. (2010). Schooling and national income: How large are the externalities? Corrected estimates. *Education Economics, 18*(4), 455-456.

4.8b 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 the expected effect size and the estimated error 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.[149] These models produced roughly 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.[150] 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.[151] 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.

$$(4.8.1) \quad Graduation_i = \alpha + \beta_1 \Delta Z_i + \beta_2 \Delta Z_i \cdot Z_{8i} + \beta_3 Z_{8i} + \delta' X_i + \xi Year_i + \epsilon_i$$

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

[149] Many, but not all, students who did not meet assessment standards in 10th grade retake exams in 11th grade.
[150] We estimate robust standard errors for the linear probability models. We also estimated logistic regression models and inferences were comparable.
[151] Low-income students are defined as ever having been eligible for free or reduced-price meals.

Exhibits 4.8.3 and 4.8.4 summarize the estimated effects of math and reading test scores on graduation status. The effects are determined by β_1 and β_2.[152] β_1 is the coefficient for the change in Z-scores. β_2 is the coefficient for an interaction term that allows the effect of test score growth to vary with the initial (8th-grade) score.

Exhibit 4.8.3

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 4.8.4

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 Monte Carlo simulations with the test score specific parameters in Exhibits 4.8.3 and 4.8.4.

The Relationship Between Gains in Student Test Scores and Labor Market Earnings. To evaluate outcomes that measure gains in student standardized test scores, the model contains a parameter and standard error to measure how a one standard deviation gain in test scores relates to a percentage increase in labor market earnings. The standard error for this input is used in Monte Carlo simulations (see Chapter 6). For these two parameters, we use regression results from Hall & Farkas (2011).[153] 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).[154] 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.[155] We enter the same parameter for all students and 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 High School Graduation and Labor Market Earnings. The model contains two types of parameters, both shown in Exhibit 4.8.5, to measure the labor market earnings effect of graduating from high school. The two types of parameters model the analytical framework established in a paper by Heckman et al. (2015).[156] One type of

[152] The effect of a change in test score is given by d(graduation)/d(ΔZ) = β1 + β2·Z8i.
[153] Hall, M. & Farkas, G. (2011). Adolescent cognitive skills attitudinal/behavioral traits and career wages. *Social Forces, 89*(4), 1261-1285.
[154] We include both the direct effect of test scores on wages as well as the indirect effect of test scores on wages through increased educational attainment.
[155] 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.
[156] Heckman et al. (2015). We use ratios of the average treatment effects as reported in Table A63 over the differences above dropouts in logged wages from that data used to create table A14 to generate our estimates. The more recent release of the paper Heckman et al. (2016). *Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking* reports similar numbers in table A40.

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. We rely on information from the Heckman et al. (2015) analysis to estimate this parameter.[157] We assume that each causal factor (percentage of the earnings difference due to the difference in education) is equal to the ratio of the average treatment effect (ATE) to the percent gain in earnings associated with reaching a particular schooling level (which was calculated using data provided by authors). Errors around the estimates are computed using the coefficients of variation calculated from the relative ATEs and the standard errors of the ATEs. These values and their 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 those in Washington aged 25 and older with some college (no degree or any degree less than a 4-year degree) and those with a 4-year degree or greater. Numbers for Juvenile Offender population estimated using information from Table 5 of the December 2016 Juvenile Justice Standardized Report *Education and Workforce Outcomes of Juvenile Justice Participants in Washington State* authored by the Education Research & Data Center at the Office of Financial Management.[158] 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 to further levels of education.

Those who continue to college incur the cost of a 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 4.8b.

4.8c Valuation of Earnings from High School Graduation

Exhibit 4.8.5
Estimates of the Causal Effect of High School Graduation on Earnings

		High school graduate (only)	Some college	4-year college graduate
Percentage of high school graduates who go on to each level of education	All students	0.25	0.38	0.37
	Low-income students	0.25	0.38	0.37
	Juvenile offenders	0.57	0.42	0.02
Percentage of observed earnings gains caused by high school graduation	Mean	0.50	0.56	0.42
	SE	0.17	0.13	0.11

The full equation for the value of a high school education is displayed in Equation 4.8.2.

$$
\begin{aligned}
(4.8.2)\quad & EarnGainHSG_y \\
&= \Big(\big(EarnHSG_y \times (1+EscHSG)^{y-age} \times (FHSG \times (1+EscFHSG)^{y-age}) \times StateAdjHSG \\
&\quad - BaselineEarn_y\big) \times \%HSG \times EarnHSGCF \\
&\quad + (EarnSomeCol_y \times (1+EscSomeCol)^{y-age} \\
&\quad \times (FSomeCol \times (1+EscFSomeCol)^{y-age}) \times StateAdjSomeCol - BaselineEarn_y) \\
&\quad + (Earn4yrDeg_y \times (1+Esc4yrDeg)^{y-age} \times (F4yrDeg \times (1+EscF4yrDeg)^{y-age}) \\
&\quad \times StateAdj4yrDeg - BaselineEarn_y) \times \%4yrDeg \times Earn4yrDegCF\Big) \times \big(IPD_{base}/IPD_{cps}\big) \\
&\quad \times (1+HCEXT)
\end{aligned}
$$

[157] Ibid.
[158] Cross, S. (2016). *Juvenile justice standardized report.*

For each year (y) throughout 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, $EarnHSG_y$ minus the earnings of a someone who did not graduate high school, $BaselineEarn_y$, a multiplied by the percentage of high school graduates who do not pursue further education, $\%HSG$, multiplied by the high school graduation causation factor, $EarnHSGCF$, multiplied by one plus the relevant real earnings escalation rate for high school graduates ($EscHSG$), raised to the number of years after program participation, multiplied by the fringe benefit rate for high school graduates ($FHSG$), multiplied by one plus the relevant fringe benefit escalation rate for all people ($EscFHSG$), raised to the number of years after program participation, multiplied by 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, $EarnSomeCol_y$, multiplied by the percentage of high school graduates who pursue some college, $\%SomeCol$, multiplied by the some college graduation causation factor, $SomeColCF$, multiplied by one plus the real earnings escalation rate for those who pursue some college ($EscSomeCol$), raised to the number of years after program participation, multiplied by the fringe benefit rate for those who pursue some college ($FSomeCol$), multiplied by one plus the relevant fringe benefit escalation rate for those who pursue some college ($EscFSomeCol$), raised to the number of years after program participation, multiplied by the ratio of state-to-national earnings for those with some college ($StateAdjSomeCol$); plus

c) The observed earnings of people with college degrees in each year, $Earn4yrDeg_y$, multiplied by the percentage of high school graduates who obtain a 4-year degree, $\%4yrDeg$, multiplied by the 4-year degree causation factor, $Earn4yrDegCF$, multiplied by one plus the real earnings escalation rate for those who obtain a 4-year degree ($Esc4yrDeg$), raised to the number of years after program participation, multiplied by the fringe benefit rate for those who obtain a 4-year degree ($F4yrDeg$), multiplied by one plus the relevant fringe benefit escalation rate for those who obtain a 4-year degree ($EscF4yrDeg$), raised to the number of years after program participation, multiplied by the ratio of state-to-national earnings for those with 4-year degrees ($StateAdj4yrDeg$); where

d) The $BaselineEarn$ is the observed earnings of people who do not graduate from high school in each year, $EarnNHSG_y$, multiplied by 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, multiplied by the fringe benefit rate of people who do not graduate from high school ($FNHSG$), multiplied by 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, multiplied by the ratio of state-to-national earnings for non-high school graduates ($StateAdjNHSG$);

e) The product is then 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 one plus the parameter for economic gain from human capital externalities, $HCEXT$.[159]

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

$$(4.8.3)\quad PVEarnGainHSG = \sum_{y=age}^{65} \frac{EarnGainHSG_y \times Units_{hsg}}{(1+Dis)^{y-age}}$$

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 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.

[159] During full 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.

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. These calculations are described in Section 4.8c, Estimating the costs of higher education and sources of revenue.

4.8e 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 4.2, and the other parameters, described in Section 4.8a, to estimate the expected gain in life cycle labor market earnings.

First, the present value of lifetime earnings is 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 Section 4.2. For each year, y, from the age of a program participant, *age*, to age 65, the modified annual CPS earnings as described in Equation 4.2.2, $ModEarnAll$, are multiplied by the degree of causation, $TSCF$, between a one standard deviation gain in student test scores and the related percentage increase in labor market earnings, multiplied by one plus the parameter for economic gain from human capital externalities, $HCEXT$.

$$(4.8.4)\quad ModEarnTest_y = \left(EarnAll_y \times (1+EscAll)^{y-age}\right) \times \left(FAll \times (1+EscFAll)^{y-age}\right) \times \left(IPD_{base}/IPD_{cps}\right) \times StateAdjAll \times TSCF \times (1+HCEXT)$$

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:

$$(4.8.5)\quad PVEarnGainTS = \sum_{y=age}^{65} \frac{ModEarnAll_y \times Units_{ts}}{(1+Dis)^{y-age}}$$

4.8f 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 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*.

$$(4.8.6)\quad PVspeced_{start} = \sum_{y=1}^{specedyears} \frac{SpecEdCostYear}{(1+Dis)^y}$$

$$(4.8.7)\quad PVspeced_{progage} = \frac{PVspeced_{start} \times \dfrac{IPD_{base}}{IPD_{speced\,costyear}}}{(1+Dis)^{start-progage}}$$

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 does not finish high school and, therefore, that the cost of this year is never incurred, this present-valued sum is multiplied by the probability of high school completion, *Hsgradprob*.

$$(4.8.8)\quad PVgraderet_{progage} = \left[\frac{EdCostYear \times \frac{IPD_{base}}{IPD_{ed\ costyear}}}{(1+Dis)^{18-progage}} \right] \times Hsgradprob$$

4.8g 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.[160]

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 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 measure 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 4.8.6). We included both IQ tests and standardized academic tests from specific program evaluations and national surveys.

Exhibit 4.8.6
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 4.8.6, 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 quadratic, cubic, logarithmic, and power models to fit a trend line to the data. A power curve provided the best combination fit (R^2=0.98) and a believable pattern of decay (Exhibit 4.8.7). 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.

[160] 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 ten; 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.

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 4.8.8 displays the adjustment factors we use in the benefit-cost model.

Exhibit 4.8.7
Estimation of Test Score Fadeout:
Meta-Analytic Results and Power Curve Model

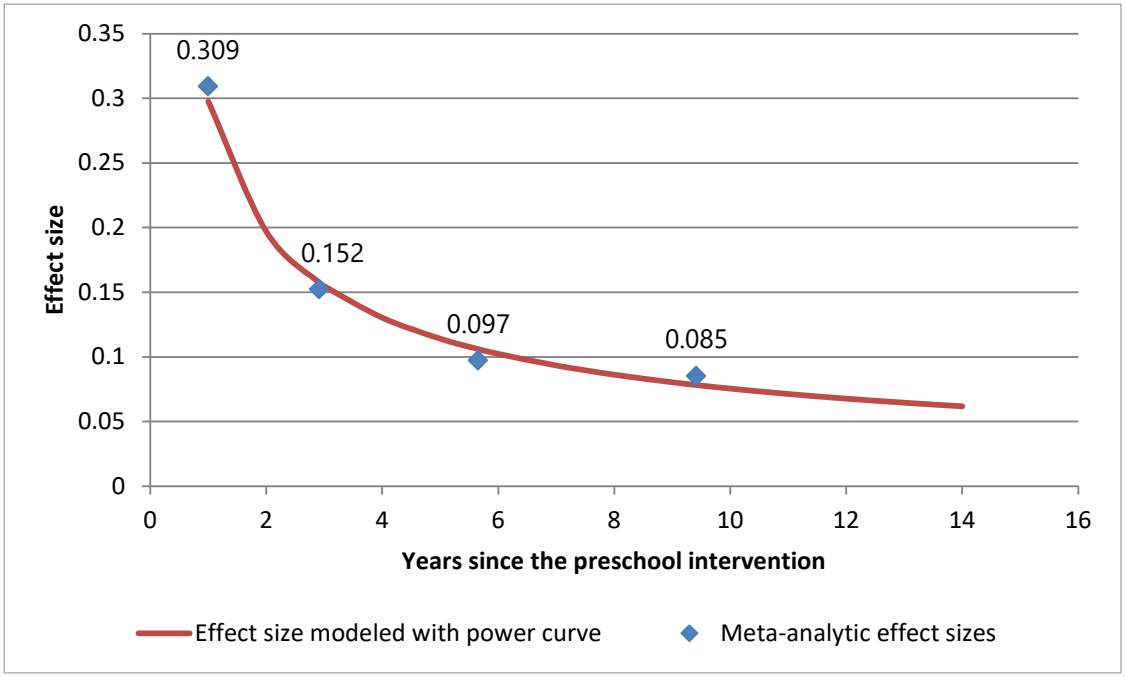


Exhibit 4.8.8

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 percentage of pre-K effect size	Fadeout multiplier: Multiply the effect size by the percent 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.

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.

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.

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.

4.9 Valuation of Higher Education Outcomes

WSIPP's benefit-cost model estimates the value of achieving certain levels of higher education through a human capital approach described in Section 4.2. The benefits of higher education programs come from increasing the probability that students obtain an education level with a higher predicted lifetime earnings trajectory than that of a high school graduate. The model moderates these gains with the financial costs (tuition, books, etc.) and opportunity costs (forgone earnings) of college attendance. We estimate the net benefit of higher education programs in two ways.

Postsecondary Attainment. The postsecondary attainment model captures the value of college enrollment, transfer, and/or graduation. We estimate the monetary benefits of higher education programs in this model by first estimating a baseline distribution of students in Washington with some college attainment, an associate's (2-year) degree, and a bachelor's (4-year) degree.[161] We then predict the change in the baseline distribution of students as a result of program participation. We monetize program impacts on one or more of the following outcomes: 2-year enrollment, 4-year enrollment, 2-year degree attainment, and 4-year degree attainment. Because these outcomes are not independent, the WSIPP model takes a comprehensive look at the relative distributions of higher education. The process is described in Section 4.9a. Section 4.9b describes how the differences in earnings gains due to the distributions are calculated, and Section 4.9c covers the calculations used to produce the costs of higher education.

Postsecondary Persistence. The persistence model captures the value of students returning to (enrolling in) any college in the years following initial enrollment. In this way, it can be thought of as a more precise measure of the returns to "some college." We estimate the monetary benefits of higher education programs in this model by first estimating a baseline percentage of students in Washington who persist to each year at either a two-year or four-year institution.[162] We then predict the change in the probability of persisting as a result of program participation. We monetize persistence as the aggregate of the program impact on one or more of the following: persistence within the first year, persistence to the second year, persistence to the third year, persistence to the fourth year,[163] and persistence to the fifth year.[164] The process is described in Section 4.9c. Section 4.9d describes how the differences in earnings gains due to the distributions are calculated, and Section 4.9e covers the calculations used to produce the costs of higher education.

4.9a Determining the Change in the Distribution of Educational Attainment Levels in the Postsecondary Attainment Model

To value postsecondary attainment we examine the lifetime earnings of people with different levels of education. The baseline distribution represents the probability a high school graduate in Washington will attain a given level of education. Changes in enrollment and graduation rates change the probabilities that students achieve higher levels of education. We monetize the differences between the baseline distribution of probabilities and the estimated distribution after applying an expected effect size from a program or intervention.

Estimating the Baseline Distribution of Educational Attainment Levels. WSIPP's benefit-cost model includes several parameters to model the likelihood that a student enrolls in and completes a degree at a 2- or 4-year institution. Exhibit 4.9.1 displays the inputs; individual inputs and their data sources are described below. The diagram in Exhibit 4.9.2 illustrates the predicted pathways of students in achieving various levels of educational attainment and the resulting baseline distribution of educational attainment levels for students in Washington. We also estimate the baseline distribution of higher educational attainment for high school students. We added this population because many of the higher education programs target K–12 students, but not all of these students will graduate from high school. We use the high school graduation rates reported in Exhibit 4.9.1 to calculate the college enrollment rate for the high school student population by multiplying the college enrollment rate for high school graduates by the high school graduation rate.

[161] We define some college attainment as enrollment in either a 2-year or 4-year institution without obtaining any degree.
[162] Because the likelihood of persistence and value of an additional year of schooling may differ at two-year versus four-year institutions, we monetize persistence for students in two-year institutions and four-year institutions separately.
[163] Only included in monetization of programs implemented at four-year institutions.
[164] Ibid.

Exhibit 4.9.1
Distribution of Higher Education Achievement

	General population		Low-income population	
	2-year college	4-year college	2-year college	4-year college
High school students				
Percentage who enroll in college	21.20%	24.88%	18.36%	13.60%
Of those who enroll, percent who graduate	31.57%	67.79%	29.34%	60.23%
High school graduates				
Percentage who enroll in college	27.14%	31.86%	27.00%	20.00%
Of those who enroll, percent who graduate	31.57%	67.79%	29.34%	60.23%
2-year college enrollees				
Percentage who graduate from 2-year institution	31.57%		29.34%	
Percentage who transfer to 4-year institution	19.18%		19.18%	
Of those who transfer, percentage who graduate from 4-year institution	56.00%		56.00%	
4-year college enrollees				
Percentage who graduate from 4-year institution		67.79%		60.23%

We use data from the State of Washington Education Research & Data Center (ERDC) to estimate the baseline percentage of high school graduates enrolling in a 2-year program, enrolling in a 4-year program, or not enrolling in higher education. Calculations are based on the 2014 enrollment percentages in ERDC's High School Feedback Reports, which measures college enrollment in the 12 months following high school graduation.[165] Estimates for low-income students are based on enrollment percentages for students receiving free- or reduced-price lunch.

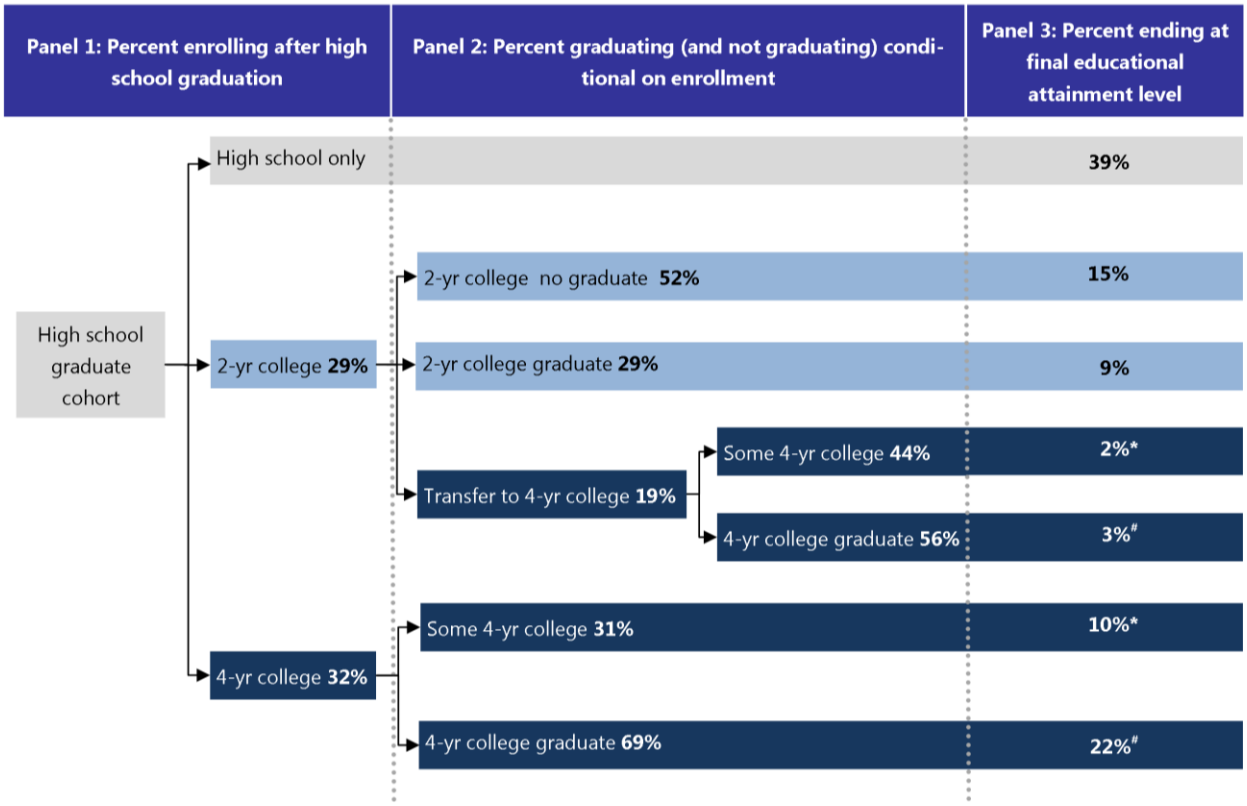
We estimate the average college graduation and transfer rates using data from the Integrated Postsecondary Education Data System (IPEDS) weighted by the number of undergraduates at the college. We calculate the proportion of students enrolled at any 4-year institution in Washington (public or private) graduating within six years using data on a cohort of students entering college in the 2010-11 academic year. We calculate the proportion of 2-year college enrollees who earn an associate's degree within three years for a cohort of students entering a Washington State 2-year institution in the 2013-14 academic year. We also calculate the proportion of students enrolled in a 2-year college who transfer to a 4-year college within three years, which we obtain using the same IPEDS data. Estimates for 4-year and 2-year low-income students are based on a subset of students who receive the federal Pell Grant, which is a grant for low-income students. We then use data from a report from the National Student Clearinghouse Research Center to determine the proportion of transfer students that graduate with a bachelor's degree.[166]

Exhibit 4.9.2 illustrates a typical Washington high school graduate's projected educational pathways for the baseline distribution. The first panel of the tree illustrates the percentage of high school graduates we estimate enroll in 2-year or 4-year colleges. The second panel of the tree shows the proportion of students that graduate and/or transfer, conditional on their initial enrollment decision. The final panel of the tree represents the final baseline distribution of high school graduates who we estimate obtain some college attainment (2- and 4-year), an associate's (2-year) degree, or a bachelor's (4-year) degree approximately six years after graduating high school.

[165] We use 2016 as it is the most current enrollment data at the time of the calculation.
[166] Shapiro, D., Dundar, A., Ziskin, M., Chiang, Y., Chen, J., Harrell, A., & Torres, V. (2013). Baccalaureate attainment: A national view of the postsecondary outcomes of students who transfer from two-year to four-year institutions. *National Student Clearinghouse Research Center.*

Exhibit 4.9.2

Higher Education Pathways Example – All High School Graduates



Final baseline distribution: Percent of students at each educational attainment level				
High school graduate only	Some college attainment		2-year degree	4-year degree
	2-year	4-year		
39%	15%	12%	9%	25%

We calculate the degree attainment by multiplying the percentage enrolling by the probability of graduating conditional on enrollment. We multiply enrollment by the percentage not graduating conditional on enrollment to estimate some college attainment. When a student can arrive at a final education level through more than one path, we sum the percentage at a final education level across all possible paths. For example, to arrive at the percentage of students with a 4-year degree we calculate the percentage with a 4-year degree through the direct path as percentage enrolling in a 4-year institution (32%) multiplied by the percentage graduating conditional on enrolling in a 4-year institution (32% x 69% = 22%). We also calculate the percentage graduating with a 4-year degree for those who start at a 2-year institution as the percentage enrolling in a 2-year institution, multiplied by the percentage of 2-year enrollees who transfer to 4-year institutions, multiplied by the percentage of transfer students who graduate (29% x 19% x 56% = 3%). We then calculate the percentage of students with a 4-year degree as the sum of these two paths (22% + 3% = 25%).

Estimating the New Distribution of Educational Attainment Levels. Our ultimate goal is to estimate the change in educational attainment due to program participation. We allow higher education programs to affect the distribution attainment in one of four ways. First, a program may change the percentage of high school graduates who attain a 2-year or 4-year degree. Second, a program can change the percentage of high school graduates who enroll at 2-year or 4-year institutions. Third, for those who are already enrolled at a 2-year or 4-year institution, the program can change the percentage of enrolled students who graduate. Finally, a program for 2-year students can change the rates at which they transfer to and/or graduate from a 4-year institution.

We apply the effect sizes estimated by each meta-analysis to the affected outcomes to determine the expected change in the baseline distribution associated with program participation.[167] For example, suppose that a program meta-analysis finds an increase in 4-year college enrollment by five percentage points but studies do not measure changes in college enrollment in 2-year institutions or overall college graduation rates. We would predict that the new rate of 4-year college enrollment from high school would be 37% (32% in the baseline distribution plus the five percentage point increase). The rate of 2-year college enrollment would remain constant at 29%, the percentage of students with a two-year degree would stay at 9%, and the new percentage of students who terminate with a high school degree would decrease by five percentage points to 34%. The conditional probabilities on the branches would remain unchanged. The rates of educational attainment for students directly enrolling in 2-year institutions remain the same (15% obtain some college attainment at a 2-year institution; 9% receive a 2-year degree; 2% transfer and attain some college attainment at a 4-year institution; and 3% transfer and attain a 4-year degree, respectively). The percentage of students who attain a 4-year degree after directly enrolling in a 4-year institution would increase to 26% (the new 4-year enrollment of 37% x 69%). Finally, the percentage of students with some college attainment who enroll directly at a 4-year institution increases to 11% (37% x 31%).[168] We monetize the change from the baseline to the new distribution as illustrated in Exhibit 4.9.3, which summarizes the above example.

Exhibit 4.9.3
Hypothetical Change in Educational Attainment Distribution

	High school graduate only	Some college attainment		2-year degree	4-year degree
		2-year	4-year		
Baseline distribution	39%	15%	12%	9%	25%
New distribution	34%	15%	13%	9%	29%
Percentage point change (Baseline—new)	-5	–	+1	–	+4

4.9b Estimating Returns to Labor Market Earnings from Changes in Postsecondary Attainment

To estimate the change in earnings as a result of postsecondary attainment, we begin with the observed earnings streams for people with varying levels of educational attainment, modified as described in Section 4.2b and illustrated in Exhibit 4.2.6. We further adjust our modified earnings streams in three ways: 1) we multiply each stream by a causal factor, 2) we remove the earnings during the time that a student is expected to spend earning that degree, and 3) we multiply the difference between modified earning streams by an externality multiplier to account for the human capital economic externalities of education as introduced in the discussion of the value of high school graduation in Section 4.9c

[167] If the increase in the probability of the affected outcome(s) is greater than the probability of the lowest educational attainment outcome then the probability of all outcomes is divided by the new base rate. For example, if a program predicts that students have a 50% chance of enrolling in a 2-year college and a 60% chance of enrolling in a 4-year college, the model assumes that students have a 45.45% chance of enrolling in a 2-year college (50/110*100%), a 54.55% chance of enrolling in a 4-year college (60/110*100%), and a 0% chance of having a high school degree only.

[168] For programs that measure enrollment *and* graduation, we estimate the new degree attainment based on the measured changes in graduation. Changes in enrollment are used to calculate the new percentage of students that obtain some college.

Our estimates of the causal increase in earnings from higher education are an extension of our high school graduation framework in Section 4.9c, which is developed from a recent paper by Heckman et al. (2015).[169] Similar to the discussion of high school graduation, we distinguish between the total difference in earnings by educational attainment and the causal difference in earnings by educational attainment using a causal factor as displayed in Exhibit 4.9.4.

Exhibit 4.9.4

Estimates of the Causal Factor of Higher Education on Earnings
(Percentage of Observed Earnings Gains Caused by Higher Education Achievement)

		Some college (2-year or 4-year)	2-year degree	4-year degree
All high school graduates	Mean	0.62	0.62	0.42
	SE	0.19	0.19	0.10
2-year college students	Mean	–	1.00	0.38
	SE		0.36	0.14
4-year college students	Mean	–	–	0.38
	SE			0.14

We assume a student has no earnings while in college, meaning we assume that the opportunity cost of college is equivalent to the total earnings for a high school graduate during the expected years in college.[170] Exhibit 4.9.5 shows the parameters we use for the expected time spent in postsecondary education.

Exhibit 4.9.5

Time Spent in Postsecondary Education

Educational pathway	Years
2-year enrollee, no transfer, no degree	1.80
2-year enrollee, transfer to 4-year, no degree	2.89
2-year enrollee, 2-year degree	3.39
2-year enrollee, transfer to 4-year, 4-year degree	4.43
4-year enrollee, no degree	2.41
4-year enrollee, 4-year degree	4.07

Note:
Years are measured in calendar years. To determine academic years spent in school, multiply calendar years by 1.33.

[169] Heckman et al. (2015). The paper by Heckman does not differentiate between levels of education below 4-year degree attainment, and it estimates the percent of the earnings difference that is causal between non-high school graduates and different levels of educational achievement. We use ratios of the average treatment effects as reported in table A63 over the percent gain in earnings associated with reaching a particular schooling level (which was calculated using data provided by authors) to generate our estimates. We assume that differences in earnings between those who attain some college and those who graduate with a 2-year degree can be wholly explained by the additional educational attainment. That is, the causal factor is one. This assumption is similar to our calculated value of 0.99 using the results presented in Marcotte et al. (2005) and is consistent with the possibility of negative selection found in the results of Brand & Xie (2010). We use the coefficient of variation of the estimate of some college to a 4-year degree to model error around this assumption. High school graduates do not have any causal increases in earnings for graduating high school nor do those enrolled in college experience gains without completing a degree. Marcotte, D.E., Bailey, T., Borkoski, C., & Kienzl, G.S. (2005). The returns of a community college education: Evidence from the National Education Longitudinal Survey. *Educational Evaluation and Policy Analysis, 27*(2), 157-175. Brand, J.E., & Xie, Y. (2010). Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *American Sociological Review, 75*(2), 273-302.

[170] We do not apply the externality multiplier to the opportunity cost for the full years a student is in school.

To calculate the time spent in school by education level we use data from the Education Longitudinal Study (ELS), which is a national survey of 10th graders in 2002 and 12th graders in 2004. We calculate the average number of months enrolled for each relevant group of students (e.g., the average months enrolled for 2-year enrollees who receive no degree and do not transfer to a 4-year institution). We use the third follow-up from 2012 and limit the analysis to students that were in 12th grade in spring 2004. Survey weights are applied to account for the complex survey design.

Gains in Earnings from Higher Education. The earnings streams are modified as described in Equation 4.2.2 of Section 4.2 to account for differences in growth rate, benefits, benefit growth, mortality, and Washington-specific factors. These adjustments create four earnings streams, $ModEarnHSG$, $ModEarnSomeCollege$, $ModEarnAssociates$, and $ModEarnBADegree$. The gains in expected earnings from higher levels of educational attainment can be simply expressed as the difference between the "baseline" stream of earnings and the stream from a higher education level, multiplied by the appropriate causal factor (from Exhibit 4.9.4), then multiplied by the economic externality factor. Interventions that change levels of post-secondary attainment often affect multiple levels of attainment, so the estimated gain in earnings resulting from a program or intervention is more complex, as shown in Equation 4.9.1.

$$
\begin{aligned}
(4.9.1)\ EarnGainHE_y &= \Big(\big((ModEarnSomeCol_y \times (1 - InSome2yrCol_y) - BasePopEarn_y \times (1 - InSchoolBase_y)\big) \\
&\times CFSome2yrfromBase \times (\%NewDistSome2yrCol - \%BaselineSome2yrCol)\Big) \\
&+ \Big(\big((ModEarnSomeCol_y \times (1 - InSome4yrCol_y) - BasePopEarn_y \times (1 - InSchoolBase_y)\big) \\
&\times CFSome4yrfromBase \times (\%NewDistSome4yrCol - \%BaselineSome4yrCol)\Big) \\
&+ \Big(\big((ModEarn2yrDegree_y \times (1 - In2yrDegree_y) - BasePopEarn_y \times (1 - InSchoolBase_y)\big) \\
&\times CF2yrDegreefromBase \times (\%NewDist2yrDegree - \%Baseline2yrDegree)\Big) \\
&+ \Big(\big((ModEarn4yrDegree_y \times (1 - In4yrDegree_y) - BasePopEarn_y \times (1 - InSchoolBase_y)\big) \\
&\times CF4yrfromBase \times (\%NewDist4yrDegree - \%Baseline4yrDegree)\Big)\Big) \times HCExt
\end{aligned}
$$

For each year (y) over the course of a person's working career, the expected earnings gain from the combination of higher education outcomes ($EarnGainHE$) is the sum of, for each adjusted attainment level (j):

a) The difference in the final distribution of attainment level from the baseline $\%NewDist(j)$ - $\%Baseline(j)$, multiplied by
b) The modified earnings stream $ModEarn(j)_y$ multiplied by 1 minus the proportion of the year spent in school in year y $In(j)_y$, subtracting
c) The modified earnings stream of the base population $BasePopEarn_y$ multiplied by 1 minus the proportion of the year spent in school in year y $InSchoolBase_y$, multiplied by
d) The causal factor determined from the two populations from Exhibit 4.9.4, used to determine what proportion of observed earnings differences is caused by higher education achievement.

To this sum, we apply a positive externality multiplier to the causal difference in earnings to reflect the benefits to society of an educated population. As earnings streams are set to 0 for the year or partial years when a student is pursuing higher education, during full years when students are attending school, we do not apply the economic gain from the human capital externality multiplier to their decreased earnings relative to non-college attendees. That is, we do not monetize negative human capital externalities. The gain in the present value of lifetime earnings from higher education attainment is estimated with this equation:

$$(4.9.2)\ PVEarnGainHE = \sum_{y=age}^{65} \frac{EarnGainHE_y}{(1 + Dis)^{y-age}}$$

4.9c Estimating Costs of Higher Education and Sources of Revenues

When an intervention increases the likelihood that an individual will attend or complete some form of higher education, there is not only a cost to implement the intervention, but a cost of increased participation in higher education that accrues to the participant and/or other funders of postsecondary education. 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 percentage of the year in school multiplied by the cost of that type of attendance (some college versus college graduate). These costs are monetized as a negative benefit and

represent a consequence (cost) of the benefits of the program (increased educational attainment) rather than a cost to implement the intervention.

Our higher education cost estimates come from our analysis of 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).[171] Exhibit 4.9.6 shows our estimates for cost and payer by type of student and education.

Exhibit 4.9.6
Higher Education Costs by Payer

	2-year institutions		Institution type unknown*		4-year institutions	
	All students	Low-income students	All students	Low-income students	All students	Low-income students
Annual cost	$10,740	$10,740	$16,312	$16,312	$22,961	$22,961
SD cost	$1,630	$1,630	$8,501	$8,501	$9,414	$9,414
Year dollars	2014	2014	2014	2014	2014	2014
Percentage paid by participant	30%	21%	47%	37%	55%	43%
Percentage paid by taxpayer	66%	75%	40%	53%	28%	41%
Federal	28%	20%	27%	21%	27%	23%
State	72%	80%	72%	79%	73%	77%
Local	0%	0%	0%	0%	0%	0%
Percentage paid by others	4%	3%	13%	10%	18%	16%

Notes:
*The costs in these columns are an average of the cost of either a 2-year or 4-year institution weighted by the number of students attending those institutions in Washington. These costs are used for estimating higher education costs for students who transfer between 2-year and 4-year institutions at an unspecified point in time, and students who attend "some college" of an unspecified type.

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.[172] 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 "4-year graduates;" the estimate using only 2-year institutions are reported as "2-year graduates," and for transfer students whose relative years at 2-year and 4-year institutions is unknown we use a per student number based on all undergraduate FTEs.

To determine the share of expenditures 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.[173] We

[171] 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).
[172] National Association of College and University Business Officers. (2002). Explaining College Costs: NACUBO's Methodology for identifying the costs of delivering undergraduate education.
[173] 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.

divide these revenues by the number of FTEs to arrive at total funding per FTE.[174] 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 the 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.[175] 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.[176] 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 percentage of all grants and scholarship dollars going to low-income students.[177] We then divide this estimate of total grant funding to low-income students by source by the percentage 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.

For each year or partial year that a person spends in higher education, we multiply the percentage of the year in school by the cost of that type of institution attended to arrive at a stream of costs for each predicted year in school. We then estimate the net present value of the stream of costs associated with attending college.[178] Using the information from the table above and the changes to the distribution of educational attainment, we estimate the change in the costs of college with this equation:

$$(4.9.3)\quad \left(\Delta CostsHE_y\right) = \Big(\Big(Cost2yr \times InSome2yrCol_y \times (\%NewDistSome2yrCol - \%BaselineSome2yrCol)\Big) + \Big(Cost4yr \times InSome4yrCol_y \times (\%NewDistSome4yrCol - \%BaselineSome4yrCol)\Big) + \Big(Cost2yr \times In2yrDegree_y \times (\%NewDist2yrDegree - \%Baseline2yrDegree)\Big) + \Big(Cost4yr \times In4yrDegree_y \times (\%NewDist4yrDegree - \%Baseline4yrDegree)\Big)\Big)$$

For each year (y) after college enrollment, the cost of college in a year $CostsHE_y$, is the sum of costs for each type of college (j):

a) The cost of a year of college institution *Cost(j)*, multiplied by,

b) The percentage of the year that the student is in school *In(j)*, multiplied by,

c) The difference in the final distribution of each level of attainment from the baseline *%NewDist(j) - %Baseline(j)*.

[174] 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.

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

[176] 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.

[177] 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 the total amount of grant and scholarship funding by income category is for all sources combined.

[178] For 2-year enrollee populations, the transfer student time spent in school and "institution type unknown" inputs are used instead of regular 4-year inputs.

Higher education costs are increased by a long-run real escalation rate in per capita inflation-adjusted higher education costs.[179] The model uses a triangular distribution around three different estimates of real higher education costs escalation (low = 0.000, modal = 0.0081, high = 0.0282). The escalation is applied beginning in the year following treatment. The present value of the costs of higher education is estimated with this equation:

$$(4.9.4)\ \Delta CostsHE = \sum_{y=age}^{65} \frac{\Delta CostsHE_y \times (1 + HEscAll)^{y-tage}}{(1 + Dis)^{y-age}}$$

4.9d Determining the Change in the Distribution of Persistence in the Postsecondary Persistence Model

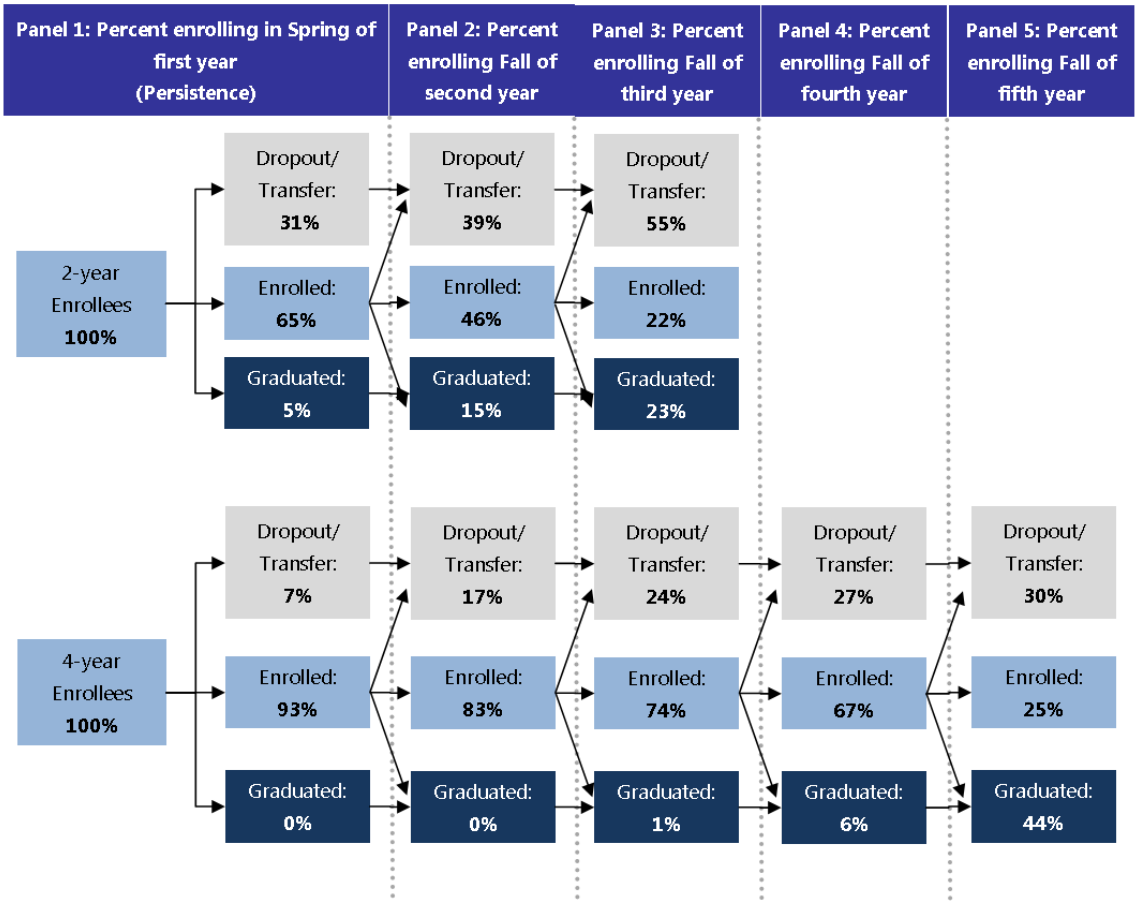
We separately value persistence, the continued year-to-year enrollment in higher education. To value persistence, we estimate the lifetime earnings of people with different years of postsecondary education. The baseline distribution represents the probability that a high school graduate in Washington will persist to a given year of education. Changes in persistence rates change the probability that students have completed a given number of years of education. We monetize the differences between the baseline distribution of probabilities and the estimated distribution after applying an expected effect size from a program or intervention. In general, persistence measures have less information than measures of postsecondary attainment. Increasing persistence in a given year may also increase the probability that a student persists to subsequent years and ultimately graduates. However, for programs that only measure persistence without measuring graduation, the change in the ultimate probability of graduation is unknown. We take a cautious approach when estimating the benefits of persistence and value persistence to a given year of postsecondary education as an increase in the probability that students have completed the previous year(s).

Estimating the Baseline Persistence Levels. WSIPP's benefit-cost model includes several parameters to model the likelihood that a student persists through a 2- or 4-year program. Exhibit 4.8.7 displays the baseline probability of persistence for students in Washington; data sources are described on the next page.

[179] The low estimate was based on the assumption that higher education costs grow at the same rate as other expenses. The middle estimate was computed based on the difference between the compound annual growth rate for the HECA education cost indexes and the IPD. The CAGR of HECA was calculated from the HECA indices from 1989 to 2014 as reported in the SHEEO technical paper. SOURCE: State Higher Education Executive Officers Association. (2014). The higher education cost adjustment: A proposed tool for assessing inflation in higher education costs. We use the differences between the CAGR of HECA, and the IPD (see section 4.11f) for the mid estimate. The high estimate was computed based on the difference between the GET program estimates of the long-term inflationary growth factor and the IPD. SOURCE: Office of the State Actuary (2017) 2017 Actuarial Valuation Report: Guaranteed Education Tuition Program.

Exhibit 4.9.7

Baseline Persistence for 2-year and 4-year Students



Final baseline distribution: Percentage of students at each persistence level					
2-year enrollees					
Initial enrollment	Persistence within first year	Persistence into second year	Persistence into third year		
100%	65%	46%	22%		
4-year enrollees					
Initial enrollment	Persistence within first year	Persistence into second year	Persistence into third year	Persistence into fourth year	Persistence into fifth year
100%	93%	83%	74%	67%	25%

We use data from the Washington State Board for Community and Technical Colleges[180] to estimate the percentage of students who enroll, graduate, or are no longer enrolled in 2-year programs. We use data from Washington State's Office of Financial Management Public Centralized Higher Education Enrollment System[181] to estimate the percentage of students who enroll, graduate, or are no longer enrolled in 4-year programs.

Estimating the New Distribution of Persistence. Our goal is to estimate the change in persistence due to program participation. When calculating the new distribution, we make the assumption that, in general, changing the probability of persisting to a given year does not change the probability in other years. For example, an observed increase in the probability of persisting to the second year is not assumed to increase the probability of persisting through the first except in special circumstances. Correspondingly, increasing persistence into the third year does not increase the probability of persisting through the fourth year. This assumption takes a cautious approach towards valuing program impacts.

[180] Calculations are based on the 2009 enrolling class. The Washington State Board for Community and Technical Colleges collects information on public community and technical colleges operating in Washington State.

[181] Calculations are based on the 2007 enrolling class. Washington State's Office of Financial Management Public Centralized Higher Education Enrollment System collects information on public 4-year institutions in Washington State.

We apply the effect sizes estimated by each meta-analysis to the persistence levels to determine the expected change in persistence associated with program participation. For example, suppose a program targeting 4-year college students increases persistence into the second year by five percentage points and persistence into the third year by three percentage points relative to the baseline but does not have any information on the impact of the program on persistence into the first year, fourth year, or fifth year. We would adjust the persistence into the second and persistence into the third year to reflect the predicted program impacts. However, the persistence within the first year, to the fourth and the fifth year would remain unchanged. The change from the baseline to the new distribution is illustrated in Exhibit 4.9.8.

Exhibit 4.9.8

Change in Baseline Persistence Rate from Hypothetical Program at 4-year Institution

Measured	Baseline persistence	Percentage point change	New calculated persistence
Persistence within first year	93%	–	93%
Persistence into second year	83%	5	88%
Persistence into third year	74%	3	77%
Persistence into fourth year	67%	–	67%
Persistence into fifth year	25%	–	25%

The only exceptions to this are when the model predicts an impossible change in persistence. For example, suppose that a program only measures persistence to the third year and predicts that more students will persist to the third year than are persisting to the second year in the baseline. In this case, the predicted persistence in the second year is impossible (greater than the observed baseline). We address this discrepancy by increasing the adjusted persistence in the second year to match the predicted third-year persistence. See Exhibit 4.9.9 for an example. Alternatively, if the model predicts that a program decreases persistence to the third year and predicts that fewer students would persist to the third year than persist to the fourth year in the baseline, then we adjust down the new predicted probability of persisting to the fourth year.

Exhibit 4.9.9

Adjustment for Impossible Program at 4-year Institution When There Is No Information

Measured	Baseline persistence	Percentage point change	Interstitial persistence	New adjusted persistence
Persistence within first year	93%	–	93%	93%
Persistence into second year	**83%**	**–**	**83%**	**89%**
Persistence into third year	74%	15	89%	89%
Persistence into fourth year	67%	–	67%	67%
Persistence into fifth year	25%	–	25%	25%

If the model predicted interstitial persistence measures are in conflict, an earlier persistence measure is given priority and serves as an upper bound for subsequent persistence measures. See Exhibit 4.9.10 for an example.

Exhibit 4.9.10

Adjustment for Impossible Program at 4-year Institution When There Is Conflicting Information

Measured	Baseline persistence	Percentage point change	Interstitial persistence	New adjusted persistence
Persistence within first year	93%	–	93%	93%
Persistence into second year	**83%**	**-4**	**79%**	**79%**
Persistence into third year	**74%**	**15**	**89%**	**79%**
Persistence into fourth year	67%	–	67%	67%
Persistence into fifth year	25%	–	25%	25%

Converting Persistence Measures to Terminal Levels of Education. Once we have determined the percentage of individuals who reach each persistence level, we calculate the implied percentage of students who stop at each level and persist no further. We use this terminal percentage to apply the appropriate predicted labor market earnings beginning at the time students have completed their education. If we applied labor market benefits to the changes in persistence levels (and not the predicted terminal level of education), we would be estimating some benefits while students are still enrolled.

We estimate the percentage of students not continuing beyond each education level (terminal percentage) from the persistence measures with the following equations.

$$(4.9.5)\quad Terminal_{i,l} = Persist_{i,l} - Persist_{i,l+1}$$

$$(4.9.6)\quad \Delta Terminal_{l} = Terminal_{n,l} - Terminal_{b,l}$$

Where:

$Persist_{i,t}$ = The baseline or new persistence percentage at year of higher education "i"
$Terminal_{i,l}$ = The baseline or new terminal percentage at year of higher education "i"
$\Delta Terminal_{l}$ = The percentage point change in the terminal percent at year of higher education "i"

Recall the example in Exhibit 4.9.8. Increasing persistence to the second year by five percentage points and persistence to the third year by three percentage points will result in the number of students stopping in the spring semester of their first year to decrease by five percentage points since these students are persisting to at least the second year. The number of students stopping in their second year is predicted to increase by two percentage points. There are five percentage points more students completing the second year of education, but the number of students who stop at the second year decreases by three percentage points since these students are continuing to the third year. This results in a two percentage point net increase in the number of students stopping at the second year. We do not know if the students who persist to their third year will continue to persist, so we make the conservative assumption that they will stop in the third year. The number of students stopping in their third year is predicted to increase by three percentage points. The change in the persistence and terminal percentages are illustrated in Exhibit 4.9.11. The changes in the terminal percentages, not the change in the persistence percentages, are used to monetize the programs.

Exhibit 4.9.11
Converting Persistence Measures to Probability of Stopping

Outcome	Baseline likelihood of persisting	Predicted percentage point change (Persistence)	New likelihood of persisting	Baseline likelihood of stopping	New Baseline likelihood of stopping	Percentage point change (Terminal)
Enroll	100.00%	0	100.00%	6.96%	6.96%	0
Persist through first year	93.04%	0	93.04%	9.62%	4.62%	-5
Persist into second year	83.42%	5	88.42%	8.70%	10.70%	2
Persist into third year	74.72%	3	77.72%	7.71%	10.71%	3
Persist into fourth year	67.01%	0	67.01%	41.78%	41.78%	0
Persist into fifth year	25.23%	0	25.23%	25.23%	25.23%	0

4.9e Estimating Returns to Labor Market Earnings from Changes in Persistence

To estimate the change in earnings as a result of persistence, we begin with the modified observed earnings streams for people with a high school degree, modified as described in Section 4.2b and illustrated in Exhibit 4.2.6. For each additional year of higher education that the student persists through, we increase the expected earnings by a persistence earnings factor. We determine the specific predicted earnings for each level of terminal education (year of enrollment in postsecondary education) by multiplying the predicted high school earnings by the persistence earnings factor.[182] The persistence earnings factor is determined by multiplying the number of years of higher education completed at each terminal education level by our estimate for the returns of an additional year of higher education.

Number of Years of Completed Higher Education. Exhibit 4.9.12 shows the parameters we use for the expected time spent in postsecondary education for each persistence (terminal education) level.

Exhibit 4.9.12
Time Spent in Higher Education

Educational pathway	Years
Persistence within first year	0.5
Persistence into second year	1
Persistence into third year	2
Persistence into fourth year	3
Persistence into fifth year	4

Estimating the Returns to an Additional Year of Higher Education. We conducted a meta-analysis to determine the expected causal increase in earnings per year that would result from an additional year of education (our persistence earnings factor). To be included, papers had to meet our normal standards for rigor (see Section 2.5 for details), analyze the returns to 2- and 4-year college education separately and control for degree receipt. By controlling for degree receipt, these results measure the returns to an additional year for students who do not complete a degree. This gives us a cautious estimate of the impact of education on earnings because it only monetizes the impact of the complete year of education. It

[182] Please note, we do not readjust the modified earnings streams to account for the differences illustrated in Exhibit 4.1.6. This is because the adjustments are made based on observational differences in earnings between education levels, not causal differences in earnings between education levels. As a result, incorporating these adjustments into the model could overestimate the benefit of persistence since some of the difference in earnings could be due to differences in the underlying characteristics of individuals who obtain different levels of education.

does not include an estimate of the increased probability of graduation, which we would also expect to increase lifetime earnings. We found two papers that met our criteria.[183] We estimated that each additional year of education at a 2-year institution would increase earnings by 6.3% over the earnings of a high school graduate. Each additional year of education at a 4-year institution would increase earnings by 6.5% over the earnings of a high school graduate. We multiply the estimated earnings increase by the number of years completed at each persistence (terminal education) level to determine the persistence earning factor, illustrated in Exhibit 4.9.13.

Exhibit 4.9.13
Estimates of the Persistence Earnings Factor of Higher Education on Earnings

Outcome		2-year degree	4-year degree
Persistence within first year	Mean	1.032	1.033
Persistence into second year	Mean	1.063	1.065
Persistence into third year	Mean	1.126	1.130
Persistence into fourth year	Mean	–	1.195
Persistence into fifth year	Mean	–	1.260

Interventions often affect multiple persistence measures, so the estimated gain in earnings in year y resulting from a program or intervention is shown in Equation 4.9.7.

$$(4.9.7)\ EarnGainPersist_y = (((ModEarnHSG_y \times \Delta Terminal_{0,y}) + (ModEarnHSG_y \times EarnPersist1_y \times \Delta\, Terminal_{1,y}) + (ModEarnHSG_y \times ModEarnPersist2_y \times \Delta Terminal_{2,y}) + (ModEarnHSG_y \times ModEarnPersist3_y \times \Delta Terminal_{3,y}) + (ModEarnHSG_y \times ModEarnPersist4_y \times \Delta Terminal_{4,y}) + (ModEarnHSG_y \times ModEarnPersist5_y \times \Delta Terminal_{5,y})) \times HCExt)$$

We set earnings streams to zero for the year or partial years when a student is pursuing higher education. During full years when students are attending school, we do not apply the economic gain from the human capital externality multiplier to their decreased earnings relative to non-college attendees. That is, we do not monetize negative human capital externalities. The gain in the present value of lifetime earnings from higher education attainment is estimated with this equation:

$$(4.9.8)\ PVEarnGainPersist = \sum_{y=age}^{65} \frac{EarnGainPersist_y}{(1+Dis)^{y-age}}$$

4.9f Estimating Costs of Persistence

We estimate the cost of persistence using the same methodology and resources outlined in Section 4.9c. Our higher education cost estimates come from our analysis of data from the Integrated Postsecondary Education Data System (IPEDS). For each year or partial year that a person spends in higher education, we multiply the percent of the year in school by the cost of the type of institution attended to arrive at a stream of costs for each predicted year in school. We then estimate the net present value of the stream of costs associated with attending college.

Using the information from Exhibit 4.9.7, and the changes to the distribution of student's predicted terminal level of education, we estimate the change in the costs of college with this equation:

$$(4.9.9)\ \Delta CostPersist_y = (Costyr_j * ((Incollege_{0,y} * \Delta Terminal_{0,y}) + (Incollege_{1,y} * \Delta Terminal_{1,y}) + (Incollege_{2,y} * \Delta Terminal_{2,y}) + (Incollege_{3,y} * \Delta\, Terminal_{3,y}) + (Incollege_{4,y} * \Delta Terminal_{4,y}) + (Incollege_{5,y} * \Delta Terminal_{5,y})))$$

[183] Marcotte et al. (2005) and Kane, T.J., & Rouse, E. (1995). Labor-market returns to two-and four-year college. *The American Economic Review, 85*(3), 600-614.

For each year (y) after college enrollment, the cost of college in a year ($CostPersist_y$) at a given institution type (j) is the sum of costs for each level of persistence (i) of:

- ✓ The cost of a year of college institution $Costyr(j)$, multiplied by ,
- ✓ An indicator of whether students with education level (i) would be enrolled in school that year $Incollege(i,y)$, multiplied by,
- ✓ The difference in the number of students who are predicted to stop at that education level $\Delta Terminal(i,y)$.

The present value of the costs of higher education is estimated with this equation:

$$(4.9.10)\ \Delta CostPersist = \sum_{y=age}^{65} \frac{\Delta CostPersist_y \text{ x } (1 + HEscAll)^{y-age}}{(1 + Dis)^{y-age}}$$

4.10 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.

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 avoiding a substantiated child abuse and neglect (CAN) case both 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 out-of-home placement. Section 4.cana describes WSIPP's calculations of CAN and OoHP prevalence in the general population and for specific subpopulations.

CAN costs are a function of four principal components. First is the expected value of public costs associated with a substantiated CAN case (e.g., child welfare system and court costs), described in Section 4.canb. Second is an estimate of the medical, mental health, and quality of life costs associated with the victim of CAN, described in section Section 4.canc and Section 4.cand. The third is the expected lifetime consequence of CAN on labor market earnings and human capital (including the higher risk of death for CAN victims compared to non-victims), described in Section 4.cane. The fourth component is made up of other long-term costs that are causally linked to the incidence of CAN; these linkages are described in Section 4.10f and further detailed in the Appendix. OoHP costs are derived from the expected value of public costs of an OoHP, 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.

Out-of-Home Placement. One component of the cost of CAN is the cost of an occurrence of an OoHP, in which a child's case is transferred to Child Welfare Services and results in a removal from the home. This subset of costs is based on the probability of an OoHP occurring within the larger CAN population. Some programs seek to prevent or directly measure an effect on OoHP. In these cases, we have additional information about the likelihood that a person experiences OoHP. We use this additional information by modeling OoHP and CAN separately.

When we value both OoHP and CAN outcomes for a single population, we use the value of our CWS system estimated through the OoHP outcome rather than the more indirect value produced from the CAN outcome. We assume that if a meta-analysis includes OoHP, we can produce a more direct estimate of the costs of removal from the home. We apply a unique set of assumptions about the spread of OoHP costs over time. When both CAN and OoHP are valued, we use the value of CPS costs as estimated through the CAN outcome.

Out-of-Home Placement Resources in the Absence of CAN. When studies in a meta-analysis report effects on out-of-home placement but do not report any measure of child abuse or neglect, we estimate the costs of OoHP as above. In addition, for all populations except for the seriously emotionally disturbed population (who are not placed due to CAN), we

assume a change in CAN equal to the magnitude of the change in OoHP. We then apply the CPS costs and direct victim costs indicated by that assumed change in CAN. The direct measure of OoHP will pick up CWS costs, so we do not compute the CWS costs indicated by the assumed change in CAN. Nor do we estimate the indirect victim costs associated with the assumed change in CAN; this is a cautious estimate given that we do not have information on whether the assumed CAN effect represents a first or subsequent event.

A Note on a Limitation 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.[184]

4.10a CAN and OoHP Prevalence

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 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 4.10.1 and 4.10.4 display the estimated prevalence rates for the analysis of child abuse and neglect and out-of-home placement in the child welfare system, respectively. Some of the rates are annual, others are cumulative; each is described in detail below.

WSIPP's CAN model begins by analyzing the national data on rates of CAN to produce estimates of the cumulative likelihood of experiencing child abuse or neglect for each age. 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. The WSIPP model combines the effect size with the base rate to calculate the estimated change in the number of avoided CAN "units" caused by the program over the lifetime of a child.

Exhibit 4.10.1 displays the following inputs, for age 1 to 18:

- ✓ The cumulative prevalence of CAN for general and low-income populations, and
- ✓ The cumulative likelihood of CAN recurrence for indicated populations.

[184] 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.

Exhibit 4.10.1
Cumulative Prevalence of Child Abuse or Neglect (CAN) by Population

Age or follow-up year	General population (by age)	Low-income population (by age)	Indicated population (by follow up year after first substantiation)
1	0.0212	0.0451	0.2124
2	0.0302	0.0635	0.3275
3	0.0389	0.0810	0.3949
4	0.0469	0.0968	0.4427
5	0.0544	0.1113	0.4797
6	0.0615	0.1247	0.5100
7	0.0681	0.1371	0.5356
8	0.0743	0.1486	0.5578
9	0.0800	0.1590	0.5774
10	0.0853	0.1687	0.5949
11	0.0903	0.1776	0.6107
12	0.0949	0.1858	0.6251
13	0.0996	0.1939	0.6384
14	0.1042	0.2020	0.6507
15	0.1088	0.2098	0.6622
16	0.1133	0.2175	0.6729
17	0.1171	0.2239	0.6830

CAN Prevalence. The likelihood of experiencing CAN varies depending on population characteristics. Furthermore, certain programs target specific populations. Most frequently in our reviews of the research, we identify three types of programs that target specific groups, which are reflected in the three columns above:

1) Broad prevention programs that serve the "general" population through universal programming.
2) Targeted prevention programs that serve families identified as high risk, often through their "low-income" status.
3) Intervention programs that aim to prevent further incidents of CAN for "indicated" children—those who already have a history of involvement with the child welfare system. These programs are "treatment" programs, as they do not prevent the first instance of CAN but instead intervene to avoid further maltreatment of prior victims of CAN.

Given the different characteristics of each of these population types, we use two basic methods to compute the estimated probability of being a victim of child abuse or neglect. First, for the general and low-income populations, we start with national data from the National Child Abuse and Neglect Data System (NCANDS), which reports the total reported CAN victims by age group per year, some of whom are repeat cases from previous maltreatment episodes.[185] NCANDS also reports the overall number of first-time victims[186] aggregated across age. To estimate the cumulative annual probability of CAN by age from this cross-sectional data, we use these two parameters to construct a synthetic cumulative probability curve, which reflects the estimated 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 4.10.1.

[185] Administration on Children, Youth and Families, (2011). *Child Maltreatment 2011* Table 3-4.
[186] Ibid, table 3-13.

$$(4.10.1)\quad CANPrev_y = \sum_{i=1}^{y}\left[\frac{Victims_i}{Pop_i} \times NewVic_i \times NewElig_i\right]$$

To compute the cumulative likelihood of CAN at age y, we use the following variables.

$Victims_i$ —the number of victims of a given age i reported by NCANDS in a year,
Pop_i—the national number of children of age i,
$NewVic_i$—the proportion of reported victims are new victims according to NCANDS (we set this parameter to one for children in their first year; it otherwise does not vary by age),
$NewElig_i$ —the proportion of children at age i who we estimate were not victims at a previous age; that is, they are eligible to be first-time victims at age i.

$$(4.10.2)\ NewElig_i = NewElig_{i-1} - CanPrev_{i-1}$$

This general prevalence curve forms the basis for our "low-income" sample as well. For the model, we estimate the increased odds of CAN for high-risk populations by taking a weighted average of the results of five studies that compared the likelihood of CAN in higher-risk populations versus lower-risk control groups (see Exhibit 4.10.2).[187]

Exhibit 4.10.2
Odds Ratios for Child Abuse and Neglect: High-Risk Populations

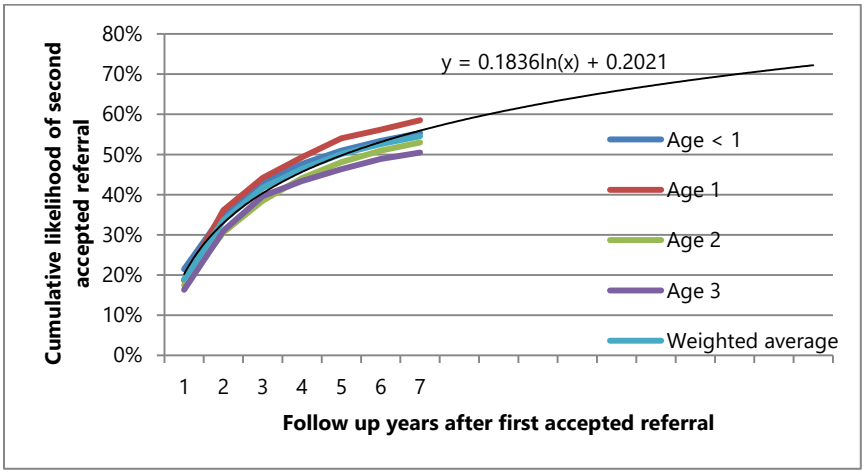
Study	Number of participants in study	Odds ratio	High-risk population
Lealman et al. (1983)	2,802	3.72	Mothers under 20, with late prenatal care, or unmarried
Murphey & Braner (2000)	29,291	2.45	Teen mothers or eligible for Medicaid
Kotch et al. (1999)	708	1.36	Receiving income support
Hussey et al.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

For the "indicated" population (children already in the child welfare system), we estimate the likelihood of *recurrence* of abuse or neglect. For this estimate of our treatment population, we use Washington State child welfare data rather than a national source; the results are displayed in Exhibit 4.10.1. We use child welfare history data from two birth cohorts in Washington State (FY 1998 and FY 2000) to estimate the proportion of those children who, after receiving one accepted referral, subsequently receive another accepted referral over time.[188] We analyze the proportion of children, first referred by age 4, who experienced a recurrence of abuse or neglect over a seven-year follow-up period, shown in Exhibit 4.10.3. We then plot a logarithmic curve with those data to predict the likelihood of a recurrence over up to 17 years after the initial incident.

[187] 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.
[188] WSIPP analysis of DSHS CAMIS data for FY 1998 and FY 2000 birth cohorts.

Exhibit 4.10.3

Predicted Likelihood of Re-referral, Based on Observations from FY 1998 and FY 2000 Birth Cohorts



OoHP Prevalence. Exhibit 4.10.4 displays the base rates of OoHP for various populations, including:

- ✓ The annual likelihood of out-of-home placement for those with CAN for general and indicated populations and
- ✓ The cumulative likelihood of out-of-home placement for the imminent risk and SED populations.

Exhibit 4.10.4

Annual and Cumulative Prevalence of Out-of-Home Placement by Population

Age or follow-up year	General population given CAN (annual, by age)	Indicated population (annual, by follow-up year since first removal)	Children at "imminent risk" of removal (cumulative, by follow-up year)	Children with serious emotional disturbance (SED) (cumulative, by follow-up year)
1	0.3439	0.3431	0.4911	0.3543
2	0.1303	0.1984	0.5682	0.4076
3	0.1127	0.1683	0.6133	0.4388
4	0.1025	0.1508	0.6453	0.4609
5	0.0952	0.1383	0.6701	0.4781
6	0.0896	0.1286	0.6903	0.4921
7	0.0849	0.1207	0.7075	0.5039
8	0.0811	0.1140	0.7223	0.5142
9	0.0777	0.1082	0.7354	0.5233
10	0.0747	0.1031	0.7471	0.5314
11	0.0720	0.0985	0.7577	0.5387
12	0.0696	0.0944	0.7674	0.5454
13	0.0674	0.0906	0.7763	0.5515
14	0.0654	0.0872	0.7846	0.5572
15	0.0635	0.0840	0.7922	0.5625
16	0.0618	0.0810	0.7994	0.5675
17	0.0601	0.0782	0.8062	0.5722

The likelihood of being placed out-of-home varies depending on population characteristics. Most frequently in our reviews of the research in which out-of-home placement has been measured, we identify four types of populations:

1. The "general" population, for which programs aim to prevent an initial event of CAN, and thereby impact the likelihood of being placed out-of-home due to maltreatment.
2. The "indicated" population, for which programs aim to prevent subsequent CAN events (and related out-of-home placement events) for children who already have a history of involvement with the child welfare system.
3. Children at "imminent risk" of placement, for which intervention programs directly target children identified as being at certain risk of removal from home in the absence of an intensive intervention.
4. Children with "serious emotional disturbance," for which intervention programs target children who are at risk of removal not for reasons of maltreatment, but due to mental health problems.

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. 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 likelihood of out-of-home placement in Washington given an accepted referral at each age.[189] Because Washington data does not allow us to capture substantiated cases, we then apply a final factor: the ratio of Washington-reported accepted referrals to estimated substantiated CAN cases.[190]

For the indicated population, we looked at all children with an accepted referral by age. We then computed the likelihood of out-of-home placement following a second accepted referral, regardless of the age of that second referral.[191]

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).[192] Our analysis resulted in a unique predicted base rate of out-of-home placement for the imminent risk population.

The last column in Exhibit 4.10.4 shows the predicted 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. Programs targeting this population and their likelihood of removal from home are rare; we used the rates of removal from the non-treated comparison groups from two studies to predict the base rate.[193]

4.10b CAN and OoHP System Cost Parameters

Estimated per-child Child Protective Services (CPS) and Child Welfare Services (CWS) system costs are displayed in Exhibit 4.10.5. 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, given an accepted referral, by the per-child cost to calculate a total expected value cost for each accepted referral.

[189] 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.

[190] 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 4.3.1).

[191] Using data from DSHS CAMIS for children born between July 1, 1997, and July 1, 2008, we looked at the proportion of those with one accepted referral by age who then received another accepted referral and were then removed from home. We then multiplied that proportion by the ratio of accepted referrals to estimated substantiated CAN cases, as described above.

[192] 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.

[193] 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.

Exhibit 4.10.5

The Estimated Average Public Cost of a Child Protective Service Case Accepted for Investigation, State of Washington (in 2016 Dollars)

$TotalSystemCost$

	Number of instances	Year of data	Per-child cost ($2016)	Probability of receiving this service	Expected cost per accepted case ($2016)
Child Protective Services (CPS)					
Referrals (children) accepted for investigation	44,246[1]	2011	$511[2]	100%	$511
Police involvement	8,053[3]	2008	$1,132[4]	18.2%	$206
Juvenile court dependency case involvement	4,864[5]	2012	$4,508[6]	19.9%	$895
In-home services (not out-of-home placement)	44,246[7]	2011	$286[7]	100%	$286
Child welfare services					
Percentage of protective custody placements that are CPS cases	96.02%[8]				
Protective custody (foster care new placements)	5,575[9]	2011	$19,271[10]	21.9%	$4,213
Adoption	1,501[1]	2011	$50,444[11]	6.1%	$3,092
Juvenile court termination case involvement	1,705[12]	2012	$4,607[6]	7.0%	$321
TOTAL: Expected present value cost of an accepted CPS case					**$9,524**
Addendum: Expected present value cost of an out-of-home placement, conditional on an out-of-home placement					$34,261
Addendum: Expected present value cost of an out-of-home placement, for a child with serious emotional disturbance (SED)[13]					$9,182
Addendum: Variation in child abuse and neglect system costs for triangle distribution					50%

Notes:

[1] WSIPP analysis of Washington State 2011 DSHS Children's Administration Data.

[2] WSIPP analysis of Washington State 2011 DSHS Children's Administration Data. Average expenditures classified for "Child Protective Services case management" on a per-child basis.

[3] Percentage (18.2%) of referrals from police sources, all states, applied to the number of total accepted referrals in 2011. From Administration on Children, Youth and Families (2015) Child Maltreatment 2015, Exhibit 2-C.

[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.

[6] WSIPP calculated an average number of hearings per case from AOC court dockets. Hearings are 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] WSIPP used the DSHS EMIS database, "Family-Focused Services" in 2011 are summed and then divided by the number of accepted referrals for a per-child estimate.

[8] Based on WSIPP analysis of DSHS Children's Administration data.

[9] Based on WSIPP analysis of DSHS Children's Administration data. The number reflects children entering foster care for reasons other than child behavior.

[10] Based on WSIPP calculation. Using DSHS Children's Administration data, WSIPP calculated the average number of days of placement in either relative placement or protective custody. The proportion of relative placements was multiplied by the calculated daily TANF rate of $1 (2016) while the proportion in protective custody was multiplied by $40 (2013), a daily rate estimate from DSHS. This average rate per day was multiplied by the average days of placement to determine the cost of placement. WSIPP added those values to the calculated cost of case management derived from DSHS Children's Administration data to create the total dollars for protective custody.

[11] WSIPP calculation of total adoption support per case, estimated from a length of adoption from DSHS data and a monthly payment rate reported Interstate Compact on the Placement of Children, Washington FY2012 Children's Administration data.

[12] Washington State Office of the Administrator of the Courts, 2012, Juvenile termination filings.

[13] 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 the home for behavior, drug, or alcohol problems between January 1, 1999, and January 1, 2005.

Sources of CAN and OoHP costs. The parameters in Exhibit 4.10.6 display the estimated proportion of system costs funding from state, local, and federal sources.

Exhibit 4.10.6
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

Notes:
[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. 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. 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).

4.10c CAN Victim Cost Parameters

Expected value victim costs are derived from calculations by Miller et al. (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 4.10.7.[194] These estimated totals are life cycle expected value costs per CAN crime; we use a procedure described in Section 4.10d to "spread out" those costs over a child's life. We use the full value victim cost when estimating the benefit of the first incident of CAN in a prevention population. When looking at a child who has already experienced an incident of CAN, we assume that quality of life costs were already incurred with the first CAN incident, and exclude those victim costs from the subsequent calculations.

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

Exhibit 4.10.7
Medical, Mental Health, and Quality of Life Costs
per Victim of Child Abuse and Neglect, 1993 Dollars
$TotalVictimCost$

	Medical and mental health costs[(1)] (1)	Quality of life costs (2)	Number of victims* (3)
Type of child abuse and neglect			
Sexual abuse	$6,327^	$94,506^	114,000
Physical abuse	$3,472^	$58,645^	308,000
Mental abuse	$2,683^	$21,099^	301,000
Serious physical neglect	$911^	$7,903^	1,236,000
Total	$1,901#	$22,948#	1,959,000
Distribution of costs by payer			
Percentage incurred by taxpayer	50%^^	0%^^	
Percentage incurred by victim	50%^^	100%^^	
Amount paid by taxpayer	**$951**	**$0**	
Amount paid by victim	**$951**	**$22,948**	
	State	**Local**	**Federal**
Victimization (taxpayer) costs##	0.500	0.000	0.500

Notes:
The source of the cost elements in this table is Miller et al. (2001).
^ *Ibid*., Table 1. We assumed 80% urban and 20% rural costs on the Miller et al. Table 1.
* The source for the total U.S. 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
# These totals are weighted average sums using the victim numbers in column (3).
^^ WSIPP assumptions.
We assume that victim costs to taxpayers are in the form of health and mental health treatment; with 50/50 FMAP split.

4.10d Procedures to Estimate CAN and OoHP System and Victim Costs

In this section, we describe how the inputs from the previous sections are used to calculate the change in expected costs caused by programs that have an impact on CAN or OoHP.

Child Abuse and Neglect Resources. Sections 4.10b and 4.10c discuss the total lifetime system and victim costs for an instance of CAN. We use these per event costs along with the prevalence rates described in Section 4.10a to estimate the costs and timing of costs as described below.

Our modeling of CAN looks at the change in the predicted amount of CAN for each year in childhood. To place the occurrence of an incident of CAN in time, we estimate the probability there is an occurrence of CAN in each year.

To estimate the timing of costs incurred within the child welfare system, we calculate the spread of lifetime costs with a "decay rate," which assumes that costs to victims are not all incurred immediately upon an event of CAN or OoHP, but rather the economic consequences continue over a number of years. We use two rates of decay: one for costs within the child welfare system, which are typically incurred all within the first eighteen years of a child's life, and one for costs to the victim, which we assume linger for a longer period.

Within the system, costs like an investigation, initial services to a family, dependency court, and so forth, occur early in a case, but child welfare services and out-of-home placements may continue for several years. We also estimate the amount of victim-related costs over time, expecting that these costs may linger much longer than system-related costs. Both estimates are described below. The proportion of total costs that occurs within a year is referred to as the *SpreadFactor* for that year.

The formulas and explanation of calculations for CAN system (Equations 4.10.3) and victim (Equation 4.10.4) are below. They are followed by the definitions and calculations for the variables used:

$$(4.10.3)\quad CAN\ System\$_b = \sum_{y=1}^{18-tage} \left[ProbOccurance_y \times SpreadFactorSystem_{t-y+1} \times (CWSCost + CPSCost) \times Unit\Delta_y\right]$$

$$(4.10.4)\quad CAN\ Victim\$_b = \sum_{y=1}^{30} \left[ProbOccurance_y \times SpreadFactorVictim_{t-y+1} \times (CAN\ Direct\ Victim\ Costs + CAN\ Indirect\ Victim\ Costs) \times Unit\Delta_y\right]$$

b—The year following the initial change in CAN or OoHP.

y—The year in the follow-up. For CAN system costs, *y* goes from age of treatment (*tage*) to 18 minus *tage*. For CAN Victim costs, *y* goes from 1 to 30.

tage—The age that a program participant is at the time when a measured program or intervention occurs.

CWSCost—Child Welfare Services costs from Exhibit 4.10.5.

CPSCost—Child Protective Services costs from Exhibit 4.10.5.

CAN Direct Victim Costs —Medical and mental health costs from Exhibit 4.10.7.

CAN Indirect Victim Costs—Quality of life costs from Exhibit 4.10.7.

$ProbOccurance_y$—Variable indicating the likelihood that a CAN event occurs in a given year. From the cumulative distributions discussed in Section 4.10a, we compute the incremental additional likelihood of CAN in each year. This hazard rate is the probability that an instance of CAN occurs in a specific year of the follow-up *y* given that an instance occurs. This probability of occurrence is adjusted so that the probability of occurrence in all years between tage and 18 is equal to 1. That is, if we assume the probability of an event happening at some point between treatment and age 18 is 1.0, $ProbOccurance_y$ estimates the likelihood of that event happening each year.

$SpreadFactor_y$.—The proportion of the lifetime costs that occur in each year *y* following an instance of CAN. We calculate our *SpreadFactorSystem* from our data in Exhibit 4.10.5. We estimate the amount of system-related costs we would expect to be incurred within the first two years of a typical CAN case (73%). 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 -0.48. That means that in the first year, 52% of the total expected costs are incurred; by the end of the second year, 73% have been incurred; 86% 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. Regardless of when we predict an incident will happen, we fit the whole predicted cost into the period from the time of the event through age 17, using the decay equation to shape the distribution of costs. Our estimated rate of decay for victim costs *SpreadFactorVictim* is -0.10 and we allow for them to continue for 30 years, 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.

$Unit\Delta_y$ —The change in the probability of experiencing CAN in year *y*.

Total CAN System and Victim Costs. Using Equations 4.2.4 and 4.2.5 adapted to the CAN topic area, we discount the sum of the change in resources and victimization costs across different types of trips and time using the following equation:

$$(4.10.5)\quad CANCost = \sum_{b=tage}^{100} \frac{(CAN\ System\$_b + CAN\ Victim\$_b)}{(1+dis)^{(b-tage+1)}}$$

As discussed in Section 4.10a, we model CAN differently for programs that prevent CAN than programs that attempt to prevent subsequent instances of CAN among a population that has previously experienced it. When an intervention is meant for a population already involved in the child welfare system, we exclude the indirect victim costs, under the assumption that those costs have already been triggered by a previous instance of CAN, and cannot be avoided.

4.10e Human Capital Outcomes—Labor Market Earnings and Deaths Attributed to CAN

Labor Market Earnings

To model the human capital outcomes affecting labor market earnings via CAN, we follow the same procedures described in depth in Section 4.5d. 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. When we combine these findings, we estimate that CAN victims earn roughly 90% of what non-victims earn.

For intervention populations, we believe that the impact of a subsequent instance of CAN on earnings is likely not as large as the impact of an initial instance of CAN. We do not have an empirically informed estimate of the magnitude of the relationship between subsequent CAN and earnings relative to initial CAN and earnings, so we apply a reduction to the magnitude to this effect size following an assumption described in Section 4.10f. We then fit distributions of expected earnings given CAN using the methodology described in Section 4.5d. Exhibit 4.10.8 shows the parameters for the fitted distributions that reflect the changes in earnings.

Exhibit 4.10.8
Labor Market Earnings 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
Expected ratio (no CAN compared to CAN)	1.118	1.053
Distribution type	Log-normal	Log-normal
Mean	-1.0761	-1.2863
Standard deviation	0.1552	0.1548
Shift	0.7777	0.7767

Deaths Attributed to CAN. Children who are victims of CAN have a higher risk of death than children who are not victims. 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.[195] We use these numbers to compute the likelihood of death by age for CAN victims (see Exhibit 4.10.9). 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 4.1d) for each age.

[195] Children's Bureau. (2015). Child abuse and neglect fatalities 2013: Statistics and interventions.

Exhibit 4.10.9

CAN attributed deaths by age, United States, 2013

Age group	Years in age group	CAN attributed deaths in U.S.	All deaths in U.S.	U.S. 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

4.10f 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 all credible studies we could locate 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 the expected effect size and the estimated error 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).

4.11 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 our methods, data sources, and estimation procedures.

The current version of WSIPP's model approaches the crime valuation question from two perspectives. First, we compute the value to taxpayers if a crime is avoided. Second, we estimate the value to would-be victims of crime, if that crime is avoided.[196] To model avoided crime costs from these two perspectives, we estimate the lifecycle costs of avoiding seven major types of crime and 11 types of costs incurred as a result of crime. In addition to computing the monetary value of avoided crime, the model estimates the number of prison beds and victimizations avoided when crime is reduced.

To monetize crime, our benefit-cost model uses four broad categories of inputs:

1) *Criminal patterns for different populations (Section 4.11a)*—These patterns serve as the basis for determining the timing and magnitude of expected costs or cost savings if a program is demonstrated to change crime outcomes.
2) *Criminal justice system probability and length of resource use (Section 4.11b)*—We estimate the likelihood that criminal justice system resources (e.g., prison or jail) will be used when a crime occurs and how long that resource will be used.
3) *Victimizations per trip (Section 4.11c)*—To capture the costs to crime victims, we estimate the total volume of reported and non-reported crime associated with a trip through the criminal justice system.

[196] 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.

4) *Criminal justice system and victim per-unit costs (Sections 4.11d and 4.11e)*—We estimate the cost of each resource within the criminal justice system and the cost of crime to victims.

This section begins by describing the methods and data sources used to estimate these four types of inputs and then turns to the computational procedures that produce the avoided costs of reduced crime.

4.11a Criminal Patterns for Different Populations

To estimate the long-run impacts of evidence-based programs on crime, WSIPP combines program effect sizes with crime information for various populations in Washington State. To establish the likelihood and timing of crime under usual circumstances, we calculate how likely it is for an average person in a specific population (e.g., individuals reentering the community from prison) to commit a crime. For the average person in each population who commits at least one crime, we estimate how many crimes they commit on average during our follow-up period, and when those crimes occur. We use 15-year recidivism trends for populations involved in the criminal justice system; for the general populations, we estimate the probability of obtaining a conviction over the life-course (50 years).

Crime Parameters. WSIPP's crime population parameters come from our analysis of our criminal history database, which combines data from the Department of Corrections and the Administrative Office of the Courts.[197] Exhibit 4.11.1 presents an example of the calculations we perform to determine the following information for each of the populations.

Cumulative Conviction Rate. We estimate the cumulative conviction rate for felony and misdemeanor crime in Washington over the 15-year (adult recidivism), 10-year (juvenile recidivism), or 50-year (life-time offense) follow-up period. We use our criminal history database to identify the first conviction for individuals during the follow-up period and compute the cumulative conviction rate using a fitted fourth-order polynomial or lognormal density distribution. These conviction rates become the base rates used to calculate the unit change of the program effect in each year of follow-up (see Section 3.2).

Total Trips through the System. We calculate the average number of "trips" through the criminal justice system during the follow-up period for each population. Each trip represents a single interaction with the criminal justice system, based on a grouping of court case numbers and date of conviction. We classify these trips into "trip types" based on the most serious offense for that trip. The mutually exclusive categories from most serious to least serious are murder, sex, robbery, assault, property, drug/other, and misdemeanor.

Trip Type Probability. For people who do commit crimes during the follow-up period, we calculate the average probability of each trip type across all trips that occurred.

Trip Timing. For those persons who incur at least one trip, we compute the average distribution of the trips in time using a probability density distribution modeled with either a fourth-order polynomial or lognormal distribution. This timing function distributes the number of trips through the system in time during the follow-up period.

[197] WSIPP's criminal history database was developed to conduct criminal justice research at the request of the legislature.

Exhibit 4.11.1

Crime Parameters from Example Population: Adult Prison (General)

Population	Number of follow-up years	Number of trips in follow-up period	Cumulative recidivism/crime over the period		Hazard rate: timing of recidivism/crime
				4^{th} order polynomial	4^{th} order polynomial
Adult prison (general)	15	4.92	Constant	0.176274	0.192420
			X	0.165020	-0.053450
			X^2	-0.024989	0.008429
			X^3	0.001725	-0.000605
			X^4	-0.000044	0.000016

Crime base population parameters	Murder	Felony sex offenses	Robbery	Aggravated assault	Felony property	Felony drug/ other	Misde-meanor
Distribution of average trips where most serious recidivism or crime offense within that trip is:	0.003	0.007	0.019	0.076	0.161	0.189	0.546

Criminal Justice-Involved 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.[198] In addition to the 15-year follow-up period (10 for juveniles), 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.
For adults, we observe recidivism patterns for 1) individuals sentenced and released from the Department of Corrections' (DOC) facilities and 2) individuals sentenced directly to DOC community supervision. We collected recidivism data on these populations who became "at-risk" for recidivism in the community during calendar years 1993-1999.
For juveniles, we observe recidivism patterns for 1) youth released from Juvenile Rehabilitation Administration (JRA) facilities and 3) youth sentenced to detention/probation/diversion/deferral through local-sanctioning courts. We collected recidivism data on these populations who became "at-risk" for recidivism in the community during 2004-2007.
We calculated separate crime distributions for each criminal justice involved population.

We further break down the general 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.[199] Additionally, based on offense of conviction we created and analyzed adult and juvenile sex offender populations and a juvenile domestic violence population.

General Population. To determine the impact of prevention programs on future crime, we calculate the probability that a person obtains a conviction over the life-course. Using WSIPP's criminal history database, we select individuals who were born between 1974 to 1977 (n=354,941) and were convicted of a felony or misdemeanor to determine how many people were convicted at age 8, age 9, age 10, and so on. The 1974 to 1977 birth cohorts allow us to use more than a single birth year and give us a long follow-up period (38 years). We extend the observed 38-year follow-up period with a probability density function to approximate a 50-year follow-up period.

In our general population calculations, the number of trips per person is the total number of trips, divided by the total unique persons observed in each cohort. The distribution of trips over time for all cohorts within the follow-up period determines trip timing, while the observed trip type determines trip probability. Our cumulative conviction rate is calculated with a series of adjustments. For each cohort, we use state population data from the Office of Financial Management to

[198] 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.
[199] 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.

abstract the number of people living in Washington State in that birth cohort year for each follow-up year. However, we adjust for whether the first trip observed for an individual is the true first trip in Washington State for that person. Since people move into and out of Washington, we need to account for the fact that many of our observed first-time individuals with a trip in the criminal justice system may have already been involved elsewhere before being convicted in Washington. We adjust the number of observed people with first trips in the criminal justice system using data from the 1997 National Longitudinal Survey of Youth (NLSY97). We compute a ratio of the first conviction compared to any conviction in a year and we apply that ratio to adjust our observed first trips in the Washington data. [200]

In addition to calculating the criminal patterns for a general population, we use this population as the basis for estimating three sub-populations, including a general population for 1) adults, 2) low-income individuals, and 3) low-income women. We use the criminological information obtained from each of these sub-populations to serve as the base rate for estimating program effects serving these populations in Washington.

General Population of Adults. Using the general population just described, analysis reveals that individuals are more likely to commit crime earlier in life (e.g., before age 30) rather than later. When estimating the effects of programs that measure crime committed by individuals in the general population greater than age 29, we use a different number of trips and crime type distribution to estimate the base likelihood of a trip occurring, as well as the distribution of trip types. We adjust our assumptions for the general population described above to account for crime that may have already have occurred. To make this adjustment, we calculate the average trips per person with a conviction and the types of trips for the later years (> 29) in our birth cohorts.

General Population of Low-Income Individuals. We also estimate criminological information for a low-income population by adjusting the general population described above using poverty and arrest data from the National Survey on Drug Use and Health.[201] 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 4.11.2). Poverty is measured as less than 200% of the federal poverty threshold. The coefficient from this model indicates that poverty is significantly related to a greater likelihood of crime ($\beta = 0.803$, $p < 0.0001$). We use the coefficient to adjust the base conviction rate (*Base*) for each year *y* over the life-course using the following equation:

$$(4.11.1)\quad AdjBase_y = \frac{\left(e^{\beta} * Base_y\right)}{\left(1 - Base_y + Base_y * e^{\beta}\right)}$$

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. We then normalize the trip crime type distribution to equal one. Our coefficients are displayed in Exhibit 4.11.2.

[200] Information is calculated using the unweighted NLSY97. Questions used are self-reported Conviction/Plead Guilty to Charges. We used a synthetic age to keep a consistent age comparison. The coefficient used is the fitted exponential factor from a regression ln(first conviction/any conviction) = B0 + B1*age. We apply this odds ratio to adjust, downward, the number of observed people with trips in Washington by year.

[201] US Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2010). *National Survey on Drug Use and Health, 2009* [Computer file]. ICPSR29621-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Exhibit 4.11.2

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

Note:
All variables were statistically significant for all models at $p < 0.001$.

Female Populations—General and Low-Income. We also estimate separate criminological information for female populations. WSIPP follows the same steps as for the general population and low-income criminological parameter estimation described above but limits the data used in the analyses to women. Exhibit 4.11.3 contains the regression results limiting our NSDUH sample only to women.

Exhibit 4.11.3

Female Population—Effect of Poverty on Arrests

	Type of arrest				
	Any	Violent	Property	Drug	Other
Intercept	-5.030	-7.076	-7.943	-7.101	-5.309
Poverty	1.062	1.223	0.986	1.191	0.980
Age 12-13	-0.242	1.239	1.775	1.124	-2.821
Age 14-15	0.886	1.658	3.007	1.316	-0.319
Age 16-17	1.199	1.522	3.187	0.872	0.515
Age 18-20	1.400	1.604	3.015	1.457	0.891
Age 21-25	1.234	1.587	2.346	1.565	0.839
Age 26-34	1.171	1.150	2.882	1.140	0.841
Black	0.025	0.584	-0.128	-1.155	-0.066
Native American	0.766	0.641	-0.322	0.655	1.003
Pacific Islander	-1.502	-0.216	-2.314	-14.514	-1.868
Asian	-1.653	-2.237	-1.842	-14.290	-1.285
Hispanic	-0.371	0.321	-0.482	-0.791	-0.608
Married	-0.848	-1.762	-0.629	-0.746	-0.844
Model Fit	0.725	0.747	0.727	0.684	0.714

Note:
All variables were statistically significant for all models at $p < 0.001$ with the exception of Pacific Islander.

4.11b Criminal Justice Probability and Length of Resource Use

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. 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. In this section, we describe how we estimate the 1) probability of each criminal justice system resource use and 2) the number of years for which the resource will be used.

Exhibit 4.11.4 below displays how criminal justice resources in Washington State are used in response to crime. We estimate the likelihood that criminal justice system resources (e.g., jail, prison) will be used when a crime occurs and the number of years the resource will be used (i.e., length of stay). For example, if an aggravated assault occurs, we estimate the chance that a person convicted of that crime will receive a prison sentence and how long the sentence will be. We updated these estimates using the most recently available Washington State data. This information is displayed in the first block of Exhibit 4.11.4. We estimate these parameters for ten types of criminal justice system resources. When possible, we calculate separate estimates for each of the seven crime types.[202]

The WSIPP model examines crime on a per-trip basis, meaning that we group convictions by distinct times where someone enters and leaves the criminal justice system. The information displayed below is on a per-trip basis, which means that it is the probability and amount of a resource that a person uses per trip (i.e., a person could have a trip for robbery that also includes consequences of a conviction for assault). The probability of jail for robbery represents the probability that anyone who has committed a robbery as the most serious crime within a trip through the system uses the jail resource. The estimates for each row in the exhibit are described below.

Juvenile Detention (with Local or State Sentence). The average length of stay for juvenile detention (9.8 days) was calculated by the Administrative Office of the Courts based on all youth whose detention stay ended in calendar year 2016.[203] The data could not be broken down by the type of sentence served (local or state sentence). The probability of resource use was based on an earlier survey of juvenile courts conducted by WSIPP. [204]

[202] Our model's counting methodology begins at the initiation of a conviction (via a trip within the criminal justice system). Thus, for police and courts, we set the probability and number of years for these resources to 1.
[203] Washington State Administrative Office of the Courts (personal communication, March 12, 2017).
[204] Burley, M., & Barnoski, R. (1997). *Washington State juvenile courts: Workloads and costs (Doc. No. 97-04-1201).* Olympia: Washington State Institute for Public Policy.

Juvenile Local Supervision. The probability of local supervision (probation) for youth in the criminal justice system and the average length of stay on probation was also estimated from a survey of juvenile courts conducted by WSIPP.[205]

Juvenile State Institution. The average length of stay in a juvenile state institution was estimated using data obtained from the Sentencing Guidelines Commission.[206]

Juvenile State Supervision. The average length of stay on juvenile parole was estimated using data obtained from the Juvenile Rehabilitation Administration.[207] We calculated the average length of stay on juvenile parole based on youth who released from an institution to parole during fiscal years 2011 and 2012.

Adult Jail, with Local Sentence. The probability of jail and the average length of stay in jail for local sentences was estimated using data obtained from the Sentencing Guidelines Commission. We calculated the length of stay based of persons sentenced during fiscal years 2011 to 2015.

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.[208]

Adult Community Supervision and Adult Post-Prison Supervision. The probability of resource use and the average length of stay for community supervision were obtained using data from the Sentencing Guidelines Commission.[209] We calculated these inputs for the two types of supervision based on persons sentenced during fiscal years 2011 to 2015.

Adult Prison. The estimates for the probability of resource use and the average length of stay in prison were calculated using sentencing data obtained from the Sentencing Guidelines Commission for Fiscal Years 2011 to 2015. The average time actually served is often shorter than the original sentence as a result of good or earned time reductions to some prison sentences.[210] Exhibit 4.11.4 shows the average prison length of stay, which is computed by multiplying the sentence length of stay by an average percentage of good/earned time reduction. The data for average sentence reductions, by crime type, were obtained from an analysis by the Washington State Department of Corrections.[211]

Technical Violations. This refers to the estimated additional length of stay in prison or jail that is experienced by those who violate the terms of their probation. In Washington, the Department of Corrections provided the length of stay in confinement, 12 days, either in prison or jail for persons who violate the terms of their community supervision. This estimate is used for those who are sentenced directly to supervision as well as for those who serve supervision after being released from prison.

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 by specific types of crimes. The last row in Exhibit 4.11.4 contains the maximum age for juvenile court jurisdiction for each type of crime. The model uses the information in Exhibit 4.11.4 as representative of the typical decisions made pursuant to current Washington State law. This information is used to determine which type of resources should be modeled in each year of an individual's modeled crime path.

[205] Ibid.
[206] Washington State Sentencing Guidelines Commission (personal communication, March 10, 2010).
[207] Washington State Juvenile Rehabilitation Administration (personal communication, April 18, 1997).
[208] Washington State Department of Corrections (personal communication, December 14, 2016).
[209] Washington State Sentencing Guidelines Commission (personal communication, April 6, 2010).
[210] The average length of a resource use is the average length for all trips within Washington, meaning that it includes the additional sentence length for subsequent trips as determined by the sentencing grid.
[211] Washington State Department of Corrections (personal communication, December 14, 2016).

Exhibit 4.11.4

Use of Crime Resources by Crime Type

Resource	Murder	Felony sex crimes	Robbery	Aggravated assault	Felony property	Felony drug/ other	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	0.02	0.40	0.24	0.54	0.59	0.63	0.00
Adult local supervision	0.81	0.84	0.79	0.68	0.26	0.62	0.00
Technical violation—Local supervision	0.31	0.31	0.31	0.31	0.31	0.31	0.00
Adult state prison	0.98	0.60	0.76	0.46	0.41	0.37	0.00
Adult post-prison supervision	0.91	0.66	0.88	0.67	0.38	0.59	0.00
Technical violation—State supervision	0.31	0.31	0.31	0.31	0.31	0.31	0.00
Number of years of resource use, if the resource is used (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, for local sentence	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Juvenile local detention, for state sentence	0.03	0.03	0.03	0.03	0.03	0.03	0.03
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	0.46	0.58	0.50	0.34	0.24	0.23	0.00
Adult jail, for prison sentence	0.80	0.48	0.46	0.46	0.38	0.32	0.00
Adult local supervision, jail sentence	1.18	2.25	1.07	1.09	1.24	1.12	0.00
Additional jail/prison time—Technical violation of local supervision	0.03	0.03	0.03	0.03	0.03	0.03	0.00
Adult state prison	16.46	4.44	3.98	2.78	1.81	1.53	0.00
Adult post-prison supervision	2.48	6.33	1.53	1.46	1.16	1.18	0.00
Additional jail/prison time—Technical violation of state supervision	0.03	0.03	0.03	0.03	0.03	0.03	0.00
Age when a juvenile is first tried in adult court							
Age	16	16	16	18	18	18	18

4.11c Estimates of Victimizations per Trip

In addition to criminal justice system costs, WSIPP estimates the number of victims and the associated costs of victimization. To account for these costs, we estimate the number of victims when a trip occurs in the criminal justice system using a combination of data from Washington State and national data sources.

When a crime occurs, multiple offenses may be processed simultaneously as a trip within the criminal justice system. We use these observed events as one basis for counting victimizations. We consider these victims associated with processed crimes as "known victims." For every trip processed by the criminal justice system, there are likely other undetected crimes that also have victims, and some of these undetected crimes are likely perpetrated by individuals processed through the criminal justice system. We consider victims of these undetected crimes "additional victims," as described below.

Known Victims per Trip. We estimate the known number of victims per trip using information about convictions from WSIPP's criminal history database. As described previously, our modeling approach is based on the unit of a trip within the criminal justice system. We classify trips hierarchically so that a trip of a particular crime type has only convictions of that crime type or a less serious type of crime associated with it. Using WSIPP's criminal history database, we calculate the average number of convictions for each trip by the most serious offense and lesser ranked offenses (i.e., a trip through the criminal justice system where the most serious conviction is for robbery may also include convictions, and victims, for assault and property crime). We assume the number of convictions as a proxy for the number of victims associated with each trip. We assume zero victims for trips where the most serious offense is drug/other or misdemeanor. See Exhibit 4.11.5 below.

Exhibit 4.11.5
Known Victims by Trip Type

		Trip type: Most serious crime associated with a trip				
		Murder	Felony sex crime	Robbery	Aggravated assault	Felony property
Victim type: Victims per trip type	Murder	1.20				
	Felony sex crime	0.01	1.64			
	Robbery	0.09	0.03	1.26		
	Aggravated assault	0.51	0.08	0.36	1.24	
	Felony property	0.05	0.03	0.20	0.20	1.71

Additional Victims per Trip. Nearly all of the effect sizes computed from programs and policies impacting crime describe official measures of criminal activity, such as convictions or arrests. Given reporting rates from the National Crime Victim Survey (NCVS), the number of crime victims using the observed victims per trip data is smaller than the "real" number of victims in Washington. These additional victims are likely not tracked or acted upon by the criminal justice system. We believe that some proportion of the victims who are unaccounted for by crimes processed through the criminal justice system are due to undetected crimes that are committed by the same perpetrators responsible for the trips captured by our analysis.

To estimate the total number of victimizations (both known and additional) per officially reported crime, WSIPP's benefit-cost model uses additional information. We calculate the total number of crimes of each type that occur in a year, calculate how many of those crimes are those observed in the criminal justice system data, and assign some proportion of the unobserved crimes to the known trips. Parameters displayed in Exhibit 4.11.6 are described below.

Exhibit 4.11.6
Estimation of Additional Victims

FBI UCR data	Victim type							
	Murder	Rape	Robbery	Aggravated assault	Burglary	Theft	Motor vehicle theft	Years of data
Number of statewide crimes reported to police	185	2,146	5,667	11,917	56,515	169,471	27,479	2011-2015
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	185	5,172	5,667	11,917	56,515	39,826	27,479	
Percentage of crime reported to police	1.0	0.307*	0.626	0.627	0.549	0.685*	0.779	2011-2015
Calculated estimate of statewide felony crimes	185	16,589	9,050	19,000	102,978	138,465	35,284	
	Murder	**Felony sex crime**	**Robbery**	**Aggravated assault**	**Felony property**			
Unreported victims	0	15,101	7,715	10,349	178,407			
Percentage of other crimes to assign to known trips	0.00	0.20	0.20	0.20	0.20			
Variation in other crimes assigned to known trips	0	0.20	0.20	0.20	0.20			
Additional victims to distribute over trips		3,020	1,543	2,070	35,681			

Note:
* These numbers rely on data from U.S. Department of Justice/Bureau of Justice Statistics. (2008). *Criminal Victimization in the United States, 2006 Statistical Tables.* National Crime Victimization Survey.

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 estimates.

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 NCVS data to estimate male victims[212] and other sexual assaults.[213] 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.[214]

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 NCVS data of thefts reported to the police to estimate this figure.[215]

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 NCVS.[216] We adjust the percentage of crimes reported to police from the NCVS for sex offenses and theft offenses differently to reflect the multiplicative adjustment to align UCR data with Washington felonies.

[212] Bureau of Justice Statistics. (2008). *Criminal victimization in the United States, 2006 statistical tables: National crime victimization survey* (Document No. NCJ 223436), Washington, DC: United States Department of Justice, Author, Table 2.
[213] Ibid., Table 1.
[214] 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.
[215] Bureau of Justice Statistics (2008), Table 100.
[216] National Crime Victimization Survey results from 2011-2015 as gathered from Bureau of Justice Statistics *Criminal victimization* series.

Percentage of Other Crimes to Assign to Known Trips. This number represents what percentage of unreported victimizations we believe are associated with observed crime trips. A value of zero would imply that those convicted of crimes are not responsible for an unobserved crime, while a value of one would imply all crimes, reported and unreported, are attributed to those convicted. To our knowledge, no research exists to date that indicates the appropriate value. We apply a "best guess" estimate of 20% for most crime types.[217]

Variance in Ratios of Other Victims per Trip. Because the additional victims per trip are estimated with considerable imprecision, we use a triangular distribution to bound the expected value in Monte Carlo simulations discussed in Chapter 7. We have chosen a lower bound of 0% and a higher bound of 40%.

The estimates in Exhibit 4.11.6 above reflect the total number of victims of each type of crime to be distributed over the trip types. We make the assumption that each trip type is only associated with crimes of that type or less serious crimes. Additional victims are distributed among those who have a trip type of an offense or a more serious type of offense based on the total number of observed victims created by each type of crime trip. The following exhibit shows these "unobserved victims" by type of crime trip and type of victim.

Exhibit 4.11.7
Additional Victims by Trip Type

		Trip type: Most serious crime associated with a trip				
		Murder	Felony sex crime	Robbery	Aggravated assault	Felony property
Victim type: Victims per trip type	Murder	0				
	Felony sex crime	0.01	2.50			
	Robbery	0.09	0.03	1.29		
	Aggravated assault	0.11	0.02	0.08	0.26	
	Felony property	0.08	0.05	0.35	0.36	2.99

In Exhibit 4.11.8, we combine the "known victims" and "additional victims" to estimate the number of victims per trip.

Exhibit 4.11.8
Total Victims by Trip Type

		Trip type: Most serious crime associated with a trip				
		Murder	Felony sex crime	Robbery	Aggravated assault	Felony property
Victim type: Victims per trip type	Murder	1.20				
	Felony sex crime	0.02	4.14			
	Robbery	0.18	0.06	2.55		
	Aggravated assault	0.61	0.10	0.43	1.51	
	Felony property	0.13	0.08	0.55	0.56	4.70

4.11d Criminal Justice System Per-Unit 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.[218] 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.

[217] As shown in Exhibit 4.11.7, we do not model additional unreported murder victims.
[218] As noted, a few average cost figures are currently used in the model when marginal cost estimates cannot be reasonably estimated.

Some short-run costs change instantly when 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. Throughout a governmental budget cycle, however, new corrections' staff are likely to be hired to reflect the change in the 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 4.11.9 and 4.11.27 display WSIPP's benefit-cost parameters for per-unit costs for the 11 sectors and seven types of crime modeled. In this section, we describe the methods used to obtain these per-unit cost estimates and the uncertainty around the estimates.

Marginal Costs and Escalation. We conducted time-series analyses of each criminal justice system resource of either panel data for Washington's 39 counties or statewide annual data. In previous iterations of WSIPP's benefit-cost model, we obtained one point estimate from one model specification to be used as the cost estimate for each criminal justice system resource. Rather than relying on the results of one regression model, we improve our cost estimates by testing a variety of model specifications for each resource.[219] We then averaged the coefficients across all the models for that resource to obtain our point estimate. This approach has two advantages. First, it allowed us to implement a variety of regression models given our understanding of the specific budget and process, including various differenced, county population-weighted, and lagged regression models so as to not rely on one model specification. Second, by averaging these coefficients, we obtained a standard deviation around each of the 11 criminal justice system estimates, which were used to estimate uncertainty for each resource-specific unit cost. We use this uncertainty when running Monte Carlo simulations in our benefit-cost model (see Chapter 7).

For each resource used, we computed an estimate of the average annual real escalation rate in costs by estimating a linear trend for each data series. From this line, we compute the predicted values for the first and last years of data and calculate the average escalation rate for the observed years, using the following formula.

$$(4.11.2)\ Rate = \ (FV/PV)^{1/N}$$

In this formula, FV is the predicted cost in the last year of data, PV is the predicted cost in the earliest year of data, and N is the number of years between the two.

[219] For each criminal justice system resource for which we estimated a time-series regression model, we ran a series of tests to address non-stationarity. Depending on the type of data (state level or panel), we used the Augmented Dickey-Fuller and the Im-Pesaran-Shin tests to test for unit roots and we used the Engle-Granger and Westerlund methods to test whether the dependent and independent variables were cointegrated. In some circumstances, we observed stationarity even after differencing, demeaning the data, or using time trends. Although stationarity is not optimal, because our estimates were reasonable compared with past analyses, we believe these results are practical estimates in the absence of any information.

Exhibit 4.11.9
Marginal Operating Costs by Crime Type

Resource	Murder	Felony sex crimes	Robbery	Aggravated assault	Felony property	Felony drug	Misde-meanor	Year of dollars	Annual real escalation rate	Per-unit cost variation
Police	1,120	1,120	1,120	1,120	1,120	1,120	1,120	2015	0.000	0.19
Juvenile local detention	51,147	51,147	51,147	51,147	51,147	51,147	51,147	2015	0.043	1.05
Juvenile local supervision	2,262	2,262	2,262	2,262	2,262	2,262	2,262	2015	0.075	0.83
Juvenile state institution	44,558	44,558	44,558	44,558	44,558	44,558	44,558	2015	0.014	0.17
Juvenile state parole	9,645	9,645	9,645	9,645	9,645	9,645	9,645	2015	0.032	0.41
Adult jail	16,776	16,776	16,776	16,776	16,776	16,776	16,776	2015	0.020	0.73
Adult local supervision	3,296	3,296	3,296	3,296	3,296	3,296	3,296	2015	0.075	0.41
Adult state prison	13,553	13,553	13,553	13,553	13,553	13,553	13,553	2015	0.001	0.10
Adult post-prison supervision	3,296	3,296	3,296	3,296	3,296	3,296	3,296	2015	0.075	0.41
Courts	152,378	18,770	9,865	4,877	201	201	201	2009	0.020	0.10

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. These cost parameters are shown in Exhibit 4.11.9.

From the Washington State Auditor, we collected local city and county police expenditure data for 1994 to 2014, all years electronically available as of winter 2016. 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 2015 dollars using the U.S. Implicit Price Deflator for Personal Consumption Expenditures from the U.S. Department of Commerce.

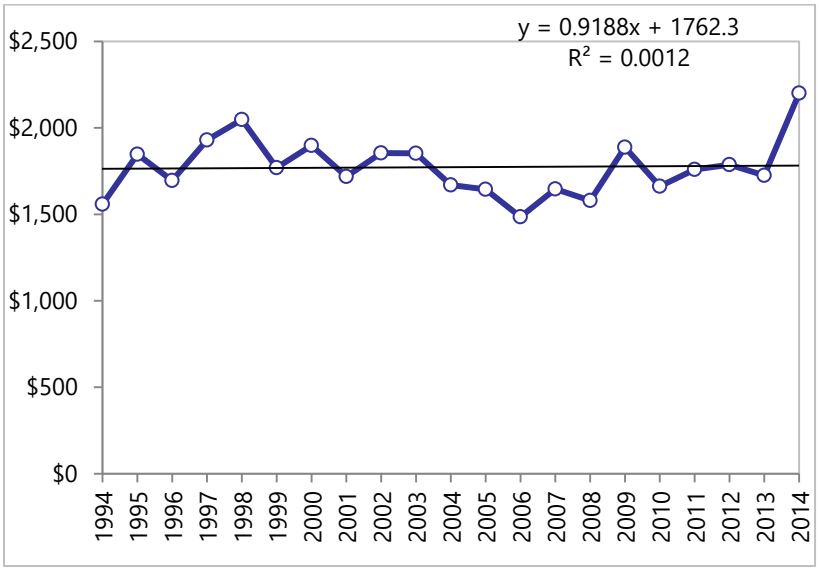
We also collected arrest information for Washington police agencies from the National Archive of Criminal Justice Data maintained by the University of Michigan.[220] Data were collected for calendar years 1994 to 2014, the earliest and latest years available as of December 2016.

The arrest data do not include the traffic operations of local police agencies. To capture this information, we obtained data from the Washington State Administrative Office of the Courts on the number of traffic infraction filings in county courts.

We aggregated the city and county expenditure data and arrest data of 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 to address newly incorporated cities where police took on responsibilities formerly assigned to county sheriffs.

[220] 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.

Exhibit 4.11.10
Average Statewide Police Costs per Arrest, 2015 Dollars
Calendar Years 1999 to 2015



Over the entire 1994 to 2014 timeframe, the average statewide cost is $1,772 per arrest, in 2015 dollars. We computed an estimate of the average annual real escalation rate in costs by estimating a linear trend (shown in Exhibit 4.11.10) for this series. From this line, we computed the predicted values for 1994 ($1,763) and 2014 ($1,782) and calculated the average escalation rate for the 21 years, using Equation 4.11.2, where FV is the 2014 estimate, PV is the 1994 estimate, and N is 20 years. We use Equation 4.11.2 to estimate an annual rate of real escalation of 0.00. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.11.9.

We tested panel data for Washington's 39 counties for 1994 to 2014. We also 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. We used statewide models but were unable to create intuitive results using disaggregated arrests. Therefore, we estimated several statewide models with total arrests. The arrest coefficients from these models were averaged to obtain the marginal cost estimate for arrests of $1,120 in 2015 dollars, as shown in Exhibit 4.11.9.

Ideally, we would be able to estimate the cost of arrest separately for each type of crime. In the future, if the data allow, we hope to examine arrests in more detail and develop an intuitive set of cost estimates, disaggregated by crime type.

Exhibit 4.11.11
Arrest Cost Regressions

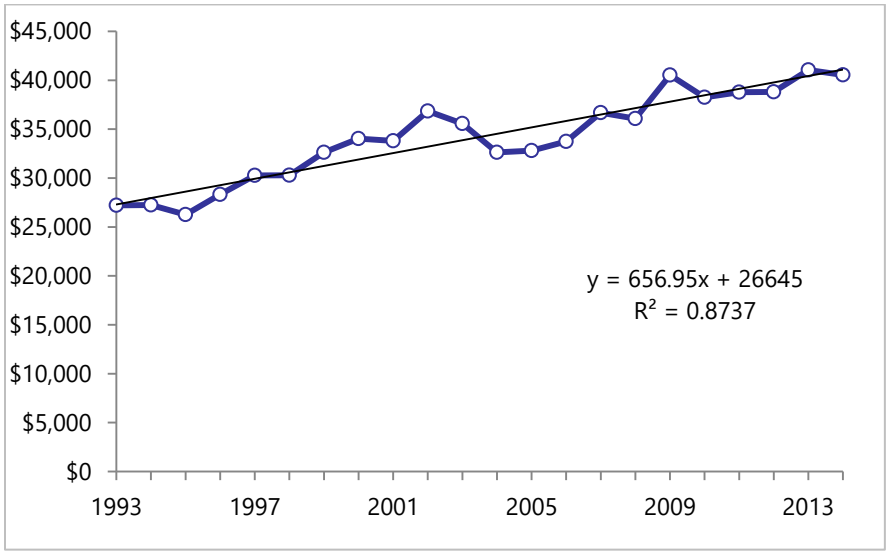
Model number	(1) Dif.StatewidePoliceCost	(2) Dif.StatewidePoliceCost
Lag.Dif.m_police_statewide	0.329	0.242
	(0.202)	(0.237)
Dif.traffic		35,136
		(21,521)
Lag.Dif.traffic		-2,735
		(23,550)
Dif.StatewideArrests	248	-51
	(407)	(448)
Lag.Dif.StatewideArrests	1,022	1,021
	(410)	(447)
Constant	3.364e + 07	3.617e + 07
	(1.146e + 07)	(1.187e + 07)
Observations	19	19
R-squared	0.408	0.521
Total	1,270	970

Local Adult Jail Per-Unit Costs. We analyze two types of users of local county-run adult jails: 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 local jail cost for both types of felons.

We collected from the Washington State Auditor local jail expenditure data for counties for 2004 to 2014, the earliest and latest years available as of winter 2016. We combined these data with information WSIPP had previously collected for the years 1993 to 2003. The Auditor's data for the expenses include all local jail expenditures (BARS code 523). These nominal annual dollar amounts were adjusted to 2015 dollars (JAILREAL) using the U.S. Implicit Price Deflator for Personal Consumption Expenditures from the U.S. Department of Commerce. Average daily jail population data (JAILADP) were obtained from the Washington Association of Sheriffs and Police Chiefs.

We computed the statewide average cost per jail average daily population (ADP) (in 2015 dollars) and plotted the results.

Exhibit 4.11.12
Average County Jail ADP Costs, 2015 Dollars
Fiscal Years 1993 to 2014



Over the entire 1993 to 2014 timeframe, the average statewide cost is $34,200 per ADP, in 2015 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 trend (shown in Exhibit 4.11.12) for this series. From this line, we computed the predicted values for 1993 ($27,302) and 2015 ($41,098) and calculated the average escalation rate, using Equation 4.11.2, where FV is the 2014 estimated cost, PV is the 1993 estimate, and N is 21 years. The annual rate of escalation is 0.020. This point estimate is included as a parameter in the crime model, as shown in
Exhibit 4.11.9.

To estimate the marginal annual operating costs of county jails, we conducted 14 panel time-series analyses of annual county-level data for jail expenditures and average jail population for each of Washington's 39 counties for calendar years 1993 to 2014. The balanced panel includes a total of 858 observations. The results of our model specifications are shown in Exhibit 4.11.13. We tested a variety of different specifications, including differencing, county population weighting (2015 population), lagging, and time periods. The jail coefficients from these models were averaged to obtain the marginal cost estimate for jail as shown in Exhibit 4.11.9.

Exhibit 4.11.13

Jail Cost Regressions (County-Year Fixed Effects)

Model number	(1) Dif.Jail Expend	(2) Jail Expend	Jail Expend	Dif.Jail Expend	Dif.Jail Expend	Dif.Jail Expend	Jail Expend	Jail Expend	Dif.Jail Expend	Jail Expend	Dif.Jail Expend	Jail Expend
Lag.Dif.Jail Expend					0.274	0.369			0.238		0.249	
					(0.0328)	(0.0304)			(0.0368)		(0.0341)	
Dif.Jail ADP	4,801			3,078	4,495	298.3			5,110		1,621	
	(1,560)			(2,000)	(1,420)	(1,690)			(1,476)		(1,696)	
Lag.Dif.Jail ADP					15,845	24,155			16,621		24,495	
					(1,418)	(1,683)			(1,467)		(1,678)	
Jail ADP		23,797	10,293				2,798	-2,125		2,783		-2,946
		(1,950)	(2,359)				(1,490)	(1,886)		(1,483)		(1,866)
Lag.Jail Expend							0.769	0.748		0.789		0.768
							(0.0159)	(0.0169)		(0.0189)		(0.0205)
Lag.Jail ADP							5,863	12,513		11,926		21,989
							(1,559)	(1,964)		(1,881)		(2,454)
TwoLag.Dif.Jail ADP									4,627		14,616	
									(1,654)		(1,955)	
TwoLag.Jail ADP										-9,003		-13,163
										(1,610)		(2,116)
Constant	343,055	621,683	1.641e+07	1.834e+06	-126,562	-418,728	-58,498	-403,480	129,710	132,497	637,366	2.246e+06
	(373,940)	(845,673)	(2.748e+06)	(966,054)	(336,053)	(816,429)	(372,112)	(1.365e+06)	(341,060)	(382,238)	(803,978)	(1.475e+06)
Observations	819	858	858	819	780	780	819	819	741	780	741	780
R-squared	0.057	0.304	0.500	0.323	0.278	0.556	0.853	0.882	0.292	0.828	0.592	0.857
Number of counties	39	39	39	39	39	39	39	39	39	39	39	39
Total	4,801	23,797	10,293	3,078	20,340	24,453	8,661	10,388	26,358	5,706	40,732	5,880

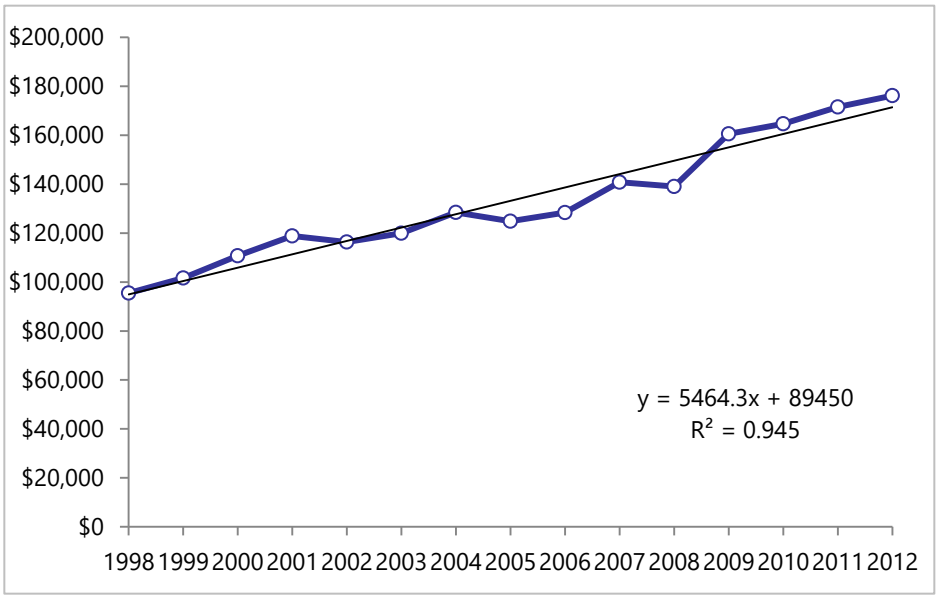
Local Juvenile Detention. For an estimate of the marginal operating cost of state juvenile offender institutions, we conduct a time-series analysis of annual data for detention expenditures and average daily admissions to juvenile detention facilities in Washington. The Washington State Auditor provided local juvenile detention operating expenditure data for counties for 2003 to 2012, the most recent year when subcategory breakouts of juvenile resource expenditures were available. We combined this information with data WSIPP had previously collected from 1998 to 2002. The Auditor's data for the expenses include the categories for residential care and custody (BARS 527.60) and juvenile facilities (BARS 527.80). Visual inspection of these historical data revealed significant problems including missing data, likely caused by inconsistent reporting, and issues with discriminating multi-jurisdictional use of detention facilities by individual counties. Additionally, discrepancies in the data categories appear to be caused by inconsistent classification practices of the expenditure categories, notably in King County. Therefore, we expand our BARS codes to include all of 527 except for 527.4, which we

consider the cost of supervision. We conduct a time-series analysis using statewide expenditures, excluding King County. These nominal annual dollar amounts were adjusted to 2015 dollars using the U.S. Implicit Price Deflator for Personal Consumption Expenditures from the U.S. Department of Commerce.

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. The average length of stay for juvenile detention is 9.8 days.[221] Using this figure, along with the actual admission data, we estimated the average daily population (ADP) of detention facilities statewide.

We computed the average costs per institutional ADP (in 2015 dollars) and plotted these data in Exhibit 4.11.14.

Exhibit 4.11.14
Average Local Juvenile Detention ADP Costs,



2015 Dollars, Fiscal Years 1998 to 2012

Over the 1998 to 2012 timeframe, the average annual cost is $133,164 per ADP, in 2015 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 trend (shown in Exhibit 4.11.14) for this series. From this line, we computed the predicted values for 1998 ($94,913) and 2012 ($171,414) and calculated the average escalation rate for the 14 years, using Equation 4.11.2, where FV is the 2012 estimated cost, PV is the 1998 estimate, and N is 14 years. The annual rate of real escalation is 0.043. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.11.9.

To estimate the marginal annual operating costs of juvenile detention, we conducted seven time-series analyses of annual statewide data for detention expenditures and average detention population for calendar years 1998 to 2012. We tested a variety of different specifications, including differencing, lagging, and time periods. The results of our model specifications are shown in Exhibit 4.11.15. The detention coefficients from these models were averaged to obtain the marginal cost estimate of $51,147 per annual ADP for juvenile detention marginal operating expenditures, in 2015 dollars, as shown in Exhibit 4.11.9.

[221] Calculated by the Administrative Office Courts based on all youth whose detention stay ended in calendar year 2016. Washington State Administrative Office of the Courts (personal communication, March 12, 2017).

Exhibit 4.11.15

Local Juvenile Detention Cost Regressions (Statewide)

Model number	(1) Dif.Local Juvenile Detention Expend	(2) Dif.Local Juvenile Detention Expend	(3) Dif.Local Juvenile Detention Expend	(4) Local Juvenile Detention Expend	(5) Local Juvenile Detention Expend	(6) Local Juvenile Detention Expend	(7) Local Juvenile Detention Expend
Lag.Dif. Local Juvenile Detention Expend		0.0335	0.00589				
		(0.273)	(0.359)				
Dif.Local Juvenile Detention ADP	71,324	75,984	55,635				
	(29,639)	(28,332)	(31,912)				
Lag.Dif.Local Juvenile Detention ADP		63,111	48,524				
		(32,859)	(40,385)				
TwoLag.Dif.Local Juvenile Detention ADP			-2,919				
			(34,636)				
Local Juvenile Detention ADP				-3,940	24,923	26,525	30,059
				(15,034)	(27,123)	(27,052)	(30,409)
Lag.Local Juvenile Detention Expend					0.596	0.565	0.406
					(0.153)	(0.155)	(0.241)
Lag.Local Juvenile Detention ADP					-11,306	1,182	6,749
					(29,852)	(32,053)	(28,230)
TwoLag.Local Juvenile Detention ADP							-27,819
							(30,861)
Year >= 2008						3,191,000	
						(3,064,000)	
Constant	2.230e+06	3.246e+06	2.093e+06	8.909e+07	2.707e+07	1.906e+07	4.774e+07
	(1.117e+06)	(1.483e+06)	(2.103e+06)	(1.014e+07)	(1.471e+07)	(1.654e+07)	(1.929e+07)
Observations	14	13	12	15	14	14	13
R-squared	0.326	0.572	0.393	0.005	0.638	0.677	0.474
Total	71,324	139,095	101,240	-3,940	13,617	27,707	8,989

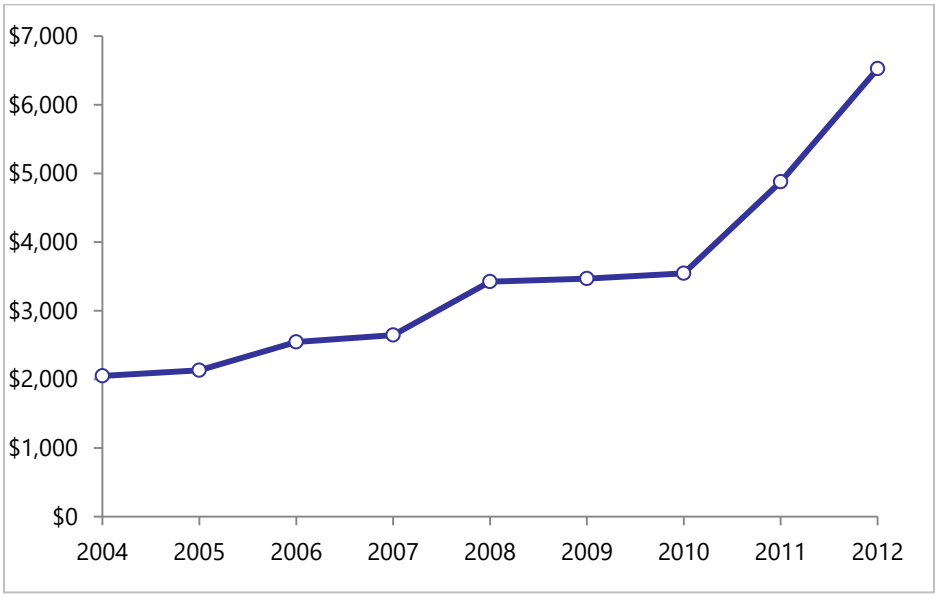
Local Juvenile Probation Per-Unit Costs. The Washington State Auditor provided local juvenile probation operating expenditure data for counties for 2003 to 2012, the most recent year when subcategory breakouts of juvenile resource expenditures were available. We combined this information with information WSIPP had previously collected from 1998 to 2002. The Auditor's data for the expenses was classified as case supervision (BARS 527.40). Unfortunately, visual inspection of these historical data revealed significant problems and gaps, likely caused by inconsistent reporting and issues determining which counties paid for which court sentences. We assume some of the discrepancies in the data categories are caused by inconsistent reporting practices, notably in King County. These nominal annual dollar amounts were adjusted to 2015 dollars using the US Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce.

From the Administrative Office of the Courts, we received the number and average term of juvenile court probation sentences for 2004 to 2014.[222] We used this information to compute an average daily population.

[222] Administrative Office of the Courts, personal communication, February 2017.

We computed the average costs per institutional ADP (in 2015 dollars) and plotted these data in Exhibit 4.11.16.

Exhibit 4.11.16
Average Local Juvenile Probation ADP Costs,
2015 Dollars, Fiscal Years 2004 to 2012



Over the entire 2004 to 2012 timeframe, the average cost is $3,468 per ADP, in 2015 dollars. Over these years we observe a spike in the inflation-adjusted costs, driven by a decline in ADP. For this reason, we used the escalation rate calculated for DOC ADP community supervision described after Exhibit 4.11.23. The annual rate of escalation is 0.075. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.11.9.

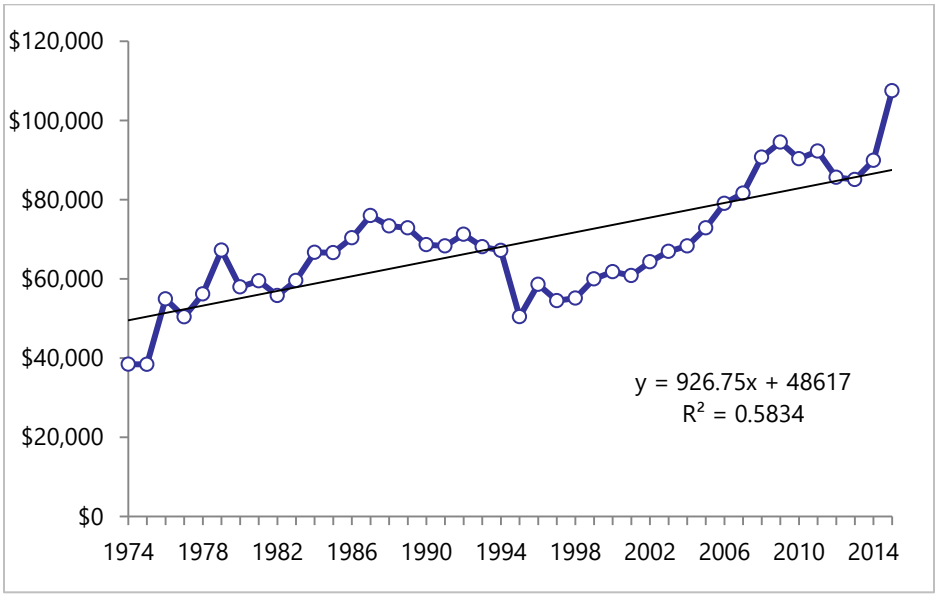
We attempted to estimate the marginal annual operating costs of juvenile probation by conducting a series of panel and time-series analyses of annual county and state-level data for probation expenditures and average daily population. After testing a variety of different specifications, including differencing and lagging, we were unable to obtain results that made intuitive sense. Instead, we used the average cost over the timeframe to estimate the marginal expenditure per average annual caseload. From our time-series analysis of the adult community supervision costs from DOC, the ratio of marginal costs to average costs was 0.652. Multiplying $3,468 by 0.652 provides a marginal cost estimate of $2,262 in 2015 dollars. This estimate is included as a parameter in the crime model, as shown in Exhibit 4.11.9.

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 State's juvenile justice agency; juvenile offenders are sentenced to JRA based on Washington's sentencing laws and practices.

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 2015. 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). We converted annual expenditure data to 2015 dollars (JRAREAL) using the U.S. Implicit Price Deflator for Personal Consumption Expenditures from the U.S. Department of Commerce. The average daily population for JRA institutions (JRAADP) series is from the Washington State Caseload Forecast Council for Fiscal Years 1997 to 2015, with data from 1974 to 1996 collected from annual reports of the Governor's Juvenile Justice Advisory Committee and 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 2015 dollars) and plotted these data in Exhibit 4.11.17.

Exhibit 4.11.17
Average JRA Institution ADP Costs, 2015 Dollars
Fiscal Years 1974 to 2015



Over the entire 1974 to 2015 timeframe, the average cost is $68,542 per ADP, in 2015 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 trend (shown in Exhibit 4.11.17) for this series. From this line, we computed the predicted values for 1974 ($49,543) and 2015 ($87,540) and calculated the average escalation rate for the 41 years, using Equation 4.11.2, where FV is the 2015 estimated cost, PV is the 1974 estimate, and N is 41 years. The annual rate of escalation is 0.014. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.11.9.

To estimate the marginal annual operating costs of JRA institutions, we conducted three time-series analyses of annual state-level data for institution expenditures and average daily population for each of calendar years 1974 to 2014. We tested a variety of different specifications, including differencing and lagging. The results of our model specifications are shown in Exhibit 4.11.18. The JRA coefficients from these models were averaged to obtain the marginal cost estimate of $44,558 for JRA institutions in 2015 dollars as shown in Exhibit 4.11.9.

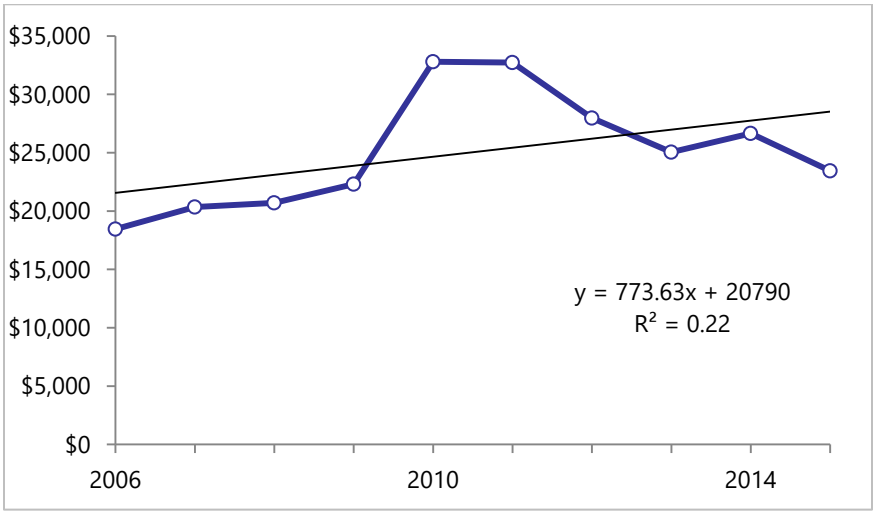
Exhibit 4.11.18
JRA Institution Cost Regressions

Model number	(1) Dif.JRA institution expenditures	(2) Dif.JRA institution expenditures	(3) Dif.JRA institution expenditures
Lag.Dif.JRA institution expenditures		-0.108	-0.103
		(0.168)	(0.145)
Dif.JRA ADP	35,687	34,731	29,972
	(7,196)	(7,781)	(6,692)
Lag.Dif.JRA ADP		12,866	11,684
		(9,460)	(8,021)
TwoLag.Dif. JRA ADP			8,735
			(6,607)
Constant	715,789	849,311	506,996
	(637,135)	(666,750)	(574,269)
Observations	41	40	39
R-squared	0.387	0.418	0.482
Total	35,687	47,597	50,391

JRA Parole Costs. To estimate the marginal operating costs of juveniles on parole after a stay at state juvenile rehabilitation facilities (JRA parole), we obtained expenditure data from the Juvenile Rehabilitation Administration's EMIS data system for fiscal years 2006 to 2015, the years following an accounting change. We converted the expenditure data to 2015 dollars (JRAParoleREAL) using the U.S. Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce. The monthly average daily population for the JRA parole (JRAParoleADP) series is from the Juvenile Rehabilitation Administration for Fiscal Years 2006 to 2015, which we adjusted to create an annual average daily population (ADP).

We computed the average costs per institutional ADP (in 2015 dollars) and plotted these data in Exhibit 4.11.19.

Exhibit 4.11.19
Average JRA Parole ADP Costs, 2015 Dollars
Fiscal Years 2006 to 2015



Over the 2006 to 2015 timeframe, the average cost is $25,045 per ADP, in 2015 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 trend (shown in Exhibit 4.11.19) for this series. From this line, we computed the predicted values for 2006 ($21,564) and 2015 ($28,526) and calculated the average escalation rate for the nine years using Equation 4.11.2,

where FV is the 2015 estimated cost, PV is the 2006 estimate, and N is nine years. The annual rate of escalation is 0.032. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.11.9.

To estimate the marginal annual operating costs of JRA parole, we conducted four time-series analyses of annual state-level data for institution expenditures and average daily population for each of calendar years 2006 to 2015. We tested a variety of different specifications, including differencing and lagging. The results of our model specifications are shown in Exhibit 4.11.20. The JRA parole coefficients from these models were averaged to obtain the marginal cost estimate for JRA annual parole of $9,645 in 2015 dollars, as shown in Exhibit 4.11.9.

Exhibit 4.11.20
JRA Parole Cost Regressions

Model number	(1) Dif.JRA parole expenditures	(2) Dif.JRA parole expenditures	(3) JRA parole expenditures	(3) JRA parole expenditures
Lag.JRA parole expenditures			0.667	0.281
			(0)	(0)
Lag.Dif.JRA parole expenditures	0.367	0.407		
	(0.460)	(0.147)		
JRA parole ADP			1,320	-672.2
			(5132)	(1463)
Lag.JRA parole ADP			6,263	741.9
			(5550)	(1801)
TwoLag.JRA parole ADP				13,443
				(1,781)
Dif.JRA parole ADP	488.3	1,994		
	(7373)	(2395)		
Lag.Dif.JRA parole ADP	-963.1	1,459		
	(7803)	(2264)		
TwoLag.Dif.JRA parole ADP		14,506		
		(2310)		
Constant	-533,980	371,198	-371,175	722,577
	(862,953)	(357,063)	(1.764e + 06)	(512,095)
Observations	8	7	9	8
R-squared	0.150	0.957	0.936	0.997
Total	-474	17,959	7,583	13,513

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.

Unlike other DOC cost estimates, the marginal cost of a prison bed is a negotiated price. DOC's budget staff estimates a marginal cost 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 4.11.21 displays the marginal costs for each legislative session. Our benefit-cost model currently uses the marginal estimate of $13,422.

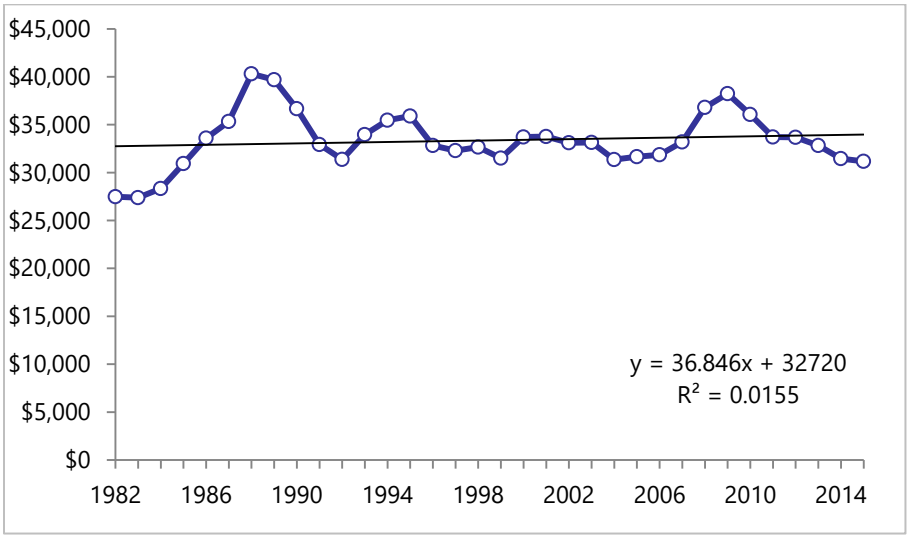
Exhibit 4.11.21
DOC Average Daily Prison Bed Marginal Cost Estimate—2014 Dollars

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 2015 dollars using the U.S. Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce. The average daily prison population (ADP) series is from the Washington State Caseload Forecast Council for fiscal years 1993 to 2015, with data for earlier years collected 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 2015 dollars) for 1982 to 2015 and plotted the results below.

Exhibit 4.11.22
Average DOC ADP Prison Costs, 2014 Dollars
Fiscal Years 1982 to 2014

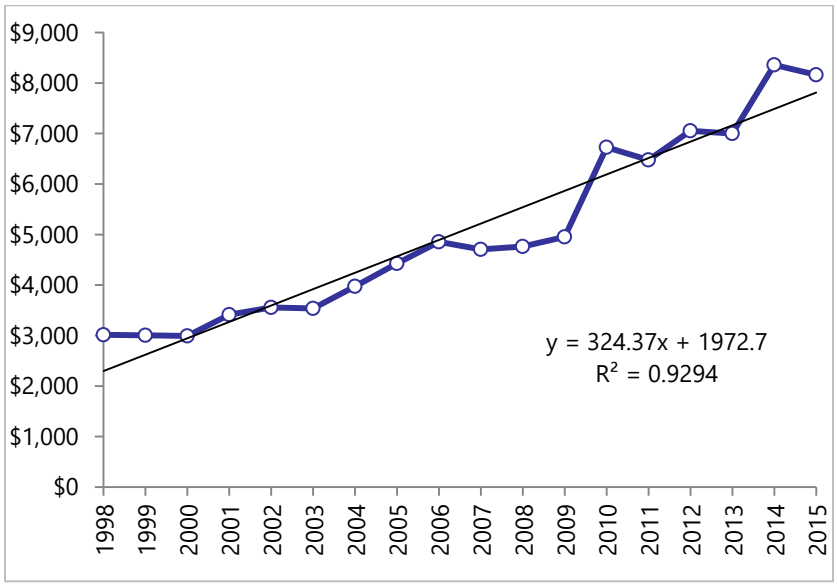


Over the 1982 to 2015 timeframe, the average cost is $33,364 per ADP, in 2015 dollars. We computed an estimate of the average annual real escalation rate in costs by estimating a linear trend (shown in Exhibit 4.11.22) for this series. From this line, we computed the predicted values for 1982 ($32,720) and 2015 ($33,972) and calculated the average escalation rate for the 33 years, using Equation 4.11.2, where FV is the 2015 estimated cost, PV is the 1982 estimate, and N is 34 years. The annual rate of escalation is 0.001. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.11.9.

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 1998 to 2015. The expenditure data were obtained from LEAP for Agency 310 (Department of Corrections) for code 300 (community supervision). Community supervision population data were obtained from the Washington Caseload Forecast Council, which maintains data back to fiscal year 1998. We calculated an annual cost per average daily community population and

converted to 2015 dollars using the aforementioned price index. The average community supervision cost over the 1998 to 2015 period is $5,054.

Exhibit 4.11.23
Average DOC ADP Community Supervision Costs,
2015 Dollars, Fiscal Years 1998 to 2015



Over the 1998 to 2015 period, there was a significant upward trend in the inflation-adjusted per-unit costs, as revealed by the linear regression line shown in Exhibit 4.11.23. To compute an estimate of the long-run growth rate in real cost per-average daily population, we calculated the predicted values from the regression line for 1998 ($2,297) and 2015 ($7,811) and calculated the annual rate of escalation for the 17 years using Equation 4.11.2 where FV is the cost estimate for 2015, PV is the estimate for 1998, and N is 17 years. The annual rate of real escalation in average costs is 0.075. This point estimate is included as a parameter in the crime model, as shown in Exhibit 4.11.9.

To estimate the marginal annual operating costs of DOC supervision, we conducted three time-series analyses of annual state-level data for supervision expenditures and average daily population for each of calendar years 1998 to 2015. We tested a variety of different specifications, including differencing and lagging. The results of our model specifications are shown in Exhibit 4.11.24. The DOC supervision coefficients from these models were averaged to obtain the marginal cost estimate for supervision, of $3,296 per annual ADP for DOC supervision expenditures, in 2015 dollars, as shown Exhibit 4.11.9.

Exhibit 4.11.24
DOC Supervision Cost Regressions

Model number	(1) Dif.DOC supervision expenditures	(2) Dif.DOC supervision expenditures	(3) Dif.DOC supervision expenditures
Lag.Dif.Supervision Expenditures		-0.0954	-0.274
		(0.26)	(0.30)
Dif.DOC Supervision ADP	1,932	1,851	2,090
	(702)	(704)	(746)
Lag.Dif.DOC Supervision ADP		1,451	1,794
		(854)	(899)
TwoLag.Dif.DOC Supervision ADP			771.2
			(846)
Constant	4.182e + 06	5.918e + 06	8.243e + 06
	(2.060e + 06)	(2.471e + 06)	(3.113e + 06)
Observations	17	16	15
R-squared	0.336	0.491	0.563
Total	1,932	3,302	4,655

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, reflecting the public costs to process cases through superior courts, which respond especially to felony crime. The cost parameters are entered into the crime model, as shown in Exhibits 4.11.9.

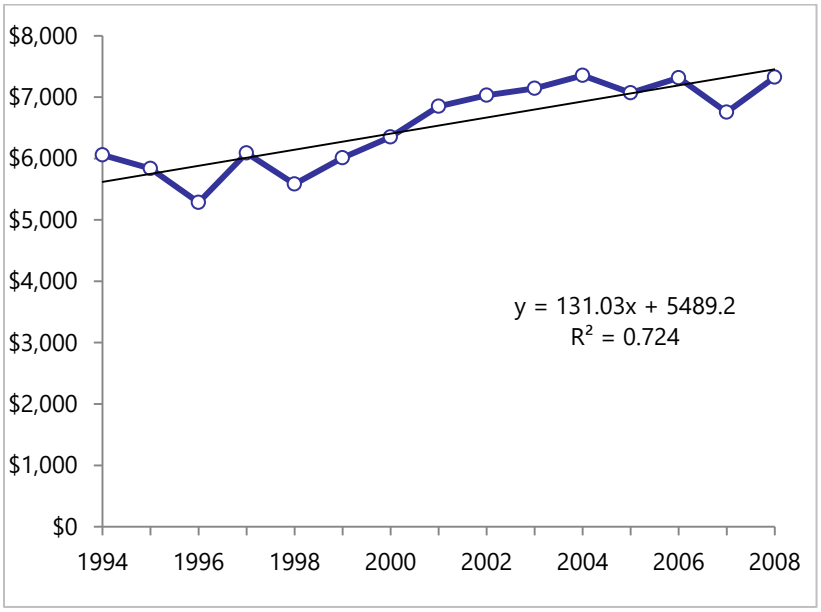
From the Washington State Auditor, we collected local county court and prosecutor expenditure data for calendar years 1994 to 2008, the earliest and latest years available as of winter 2010.[223] The Auditor's data for the expenses include all local court and prosecutor expenditures (BARS code 512 for courts and BARS code 515 for prosecutors). The court data include the costs of administration (BARS 512.10), superior courts (BARS 512.20), and county clerks (BARS 512.30). For court expenditure data, we excluded district courts (BARS 512.40), since they do not process felony cases (the main subject of interest in our benefit-cost analysis) and expenditures for law library (BARS 512.70) and indigent defense (BARS 512.80); this latter category was excluded because the data were not available for the entire time frame under review. The prosecutor data include costs for administration-legal (515.10) and legal services (515.2). For prosecutor offices, we excluded facilities-legal services (515.50), consumer affairs-legal services (515.60), crime victim and witness program-legal (515.70), and child support enforcement-legal services (515.80). All nominal annual dollar amounts were adjusted to 2009 dollars using the US Implicit Price Deflator for Personal Consumption Expenditures from the US Department of Commerce.

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.

[223] In 2016 we also retrieved more recent data. Visual inspection of these historical data revealed significant problems including missing data, likely caused by inconsistent reporting. We rely on our previous estimates and data collection efforts of information from 1999 to 2008.

Exhibit 4.11.25
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 trend (shown in Exhibit 4.11.25) 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 4.11.2, 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 4.11.9.

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 2014. However, we were unable to obtain results that made intuitive sense across all seven crime categories. Until we can improve the data or model specifications, we rely on our previously estimated marginal operating costs of court, relying on data from 1999 to 2008.

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. We tested a variety of different specifications, including differencing and lagging.[224] The results of our model specification produced five crime-specific cost estimates shown in Exhibit 4.11.9.

[224] Our preferred model was a first-difference model where we included lags of each of the violent felony conviction variables along with a variable for all other convictions, as well as county and time fixed effects. We also included a lagged dependent variable. This model produced coefficients for the violent conviction variables that made the most intuitive sense.

Exhibit 4.11.26
DOC Supervision Cost Regressions

Model number	(1) Dif.Court expenditures
Lag.Dif.Court expenditures	-0.113 (0.169)
Lag.Dif.MurderConviction	152,377.9 (125,366.9)
Lag.Dif.SexCrimeConviction	18,770.28 (11,395.58)
Lag.Dif.RobberyConviction	9,865.480 (29,782.45)
Lag.Dif.AssaultConviction	4,876.710 (9,512.385)
Lag.Dif.NonViolentFelonyConviction	200.5611 (1,503.985)
Constant	15,8006.5 (86,235.19)
Observations	10
R-squared	0.209
Number of counties	39

Capital Costs. WSIPP includes the capital allocation of detention facilities in our criminal justice system marginal cost estimates. In our crime model, the total capital cost per bed is converted to an annualized capital payment, 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, as given by the following equation:

$$(4.11.3) \quad PMT = \frac{iPV}{1-(1+i)^{-n}}$$

When computing the costs of a criminal justice system resource, WSIPP combines the marginal and capital costs for the resource and applies the escalation costs listed in Exhibit 4.11.9. When conducting Monte Carlo analysis, WSIPP draws from a triangular cost distribution using the parameters listed in Exhibit 4.11.9.

Jail Capital Costs. Local adult jail capital costs for new beds were based on the experience of the SCORE facility.[225] We used the budgeted $97 million over the 802 beds, resulting in a $120,948 capital cost in 2009 dollars per county jail bed.

Local Detention Capital Costs. Per-bed capital costs for a new detention facility would run $200,000 per bed in 2009 dollars.[226]

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.

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 (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 4.11.27.

[225] 2012 Municipal Excellence Awards Entry Form.
[226] Capital costs for a typical new local juvenile detention facility were estimated from personal communication with Washington's Juvenile Rehabilitation Administration staff.

Exhibit 4.11.27
Capital Costs for Crime Resources

Resource	Capital cost per unit	Year of dollars	Finance years	Per year capital cost calculation
Police	n/a	n/a	n/a	n/a
Courts	n/a	n/a	n/a	n/a
Juvenile local detention	200,000	2009	25	15,997
Juvenile local supervision	n/a	n/a	n/a	n/a
Juvenile state institution	150,000	2009	25	11,998
Juvenile state supervision	n/a	n/a	n/a	n/a
Adult jail	120,948	2009	25	9,674
Adult local supervision	n/a	n/a	n/a	n/a
Adult state prison	113,339	2007	25	9,329
Adult post-prison supervision	n/a	n/a	n/a	n/a

Criminal Justice Costs by Funding Source. Exhibit 4.11.28 shows the breakouts and sources of criminal justice costs for Washington State.

Exhibit 4.11.28
Proportional of Marginal Criminal Justice Costs by Funding Source

	Operating			Capital		
	State	Local	Federal	State	Local	Federal
Police^	14%	86%	0%	n/a	n/a	n/a
Courts & prosecutors^	16%	84%	0%	n/a	n/a	n/a
Juvenile local detention	15%*	85%	0%	0%#	100%	0%
Juvenile local supervision	15%*	85%	0%	n/a	n/a	n/a
Juvenile state institution^^	100%	0%	0%	100%	0%	0%
Juvenile state supervision^^	100%	0%	0%	n/a	n/a	n/a
Adult jail**	25%	75%	0%	0%^^	100%	0%
Adult local supervision^^	100%	0%	0%	n/a	n/a	n/a
Adult state prison^^	100%	0%	0%	100%	0%	0%
Adult post-prison supervision^^	100%	0%	0%	100%	0%	0%

Notes:
^ Justice Expenditure and Employment Extracts, 2012—Preliminary, Tracey Kyckelhahn, Ph.D., July 1, 2013. NCJ 242544, Table 4: Justice system expenditure by character, state and type of government, fiscal 2012. Direct current Police Protection expenditures for state and local governments for Washington State.
* Calculated using local operating expenditures costs and state pass-through funds for 2011. Operating costs come 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). Detention and Correction (BARS account: 527). 2011 State expenditures from BARS. 2011 state juvenile court pass-through funding comes from personal communication with Cory Redman, DSHS, April 25, 2017.
# WSIPP assumes capital costs for all local juvenile and adult resources are 100% locally funded.
^^ WSIPP assumes all state funded.
** WSIPP assumption.

4.11e 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.[227] These crime victim costs build on and modify the previous work prepared for the U.S. Department of Justice by Miller et al. (1996).[228]

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 estimate 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 4.11c of this chapter for a description of WSIPP's data sources for counting convictions.

The Cohen & Piquero 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 & Piquero 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 & Piquero studies 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 4.11.29.
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.

[227] 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.
[228] 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.

Exhibit 4.11.29
Victim Costs

Resource	Murder	Felony sex crimes	Robbery	Aggravated assault	Felony property	Year of dollars (of data)
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

4.11f 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, *r*, as described in Exhibits 4.11.9 and 4.11.27, we estimate costs avoided using the following equation:

$$(4.11.4)\quad CjsResource\$_{rb} = \sum_{c=1}^{C}\sum_{f=1}^{F}\left[CjsEvent_{bcf} \times CjsResourcePr_{rc} \times CjsResourceCost_{rc}(1+CjsResourceCostEsc_{r})^{(b-1)} \times TotalTrips \times TripTiming_{f} \times TripTypePr_{c} \times Unit\Delta_{f}\right]$$

We also count Average Daily Population prison beds avoided. We do this using Equation 4.11.4 above however; we do not multiply by the $CjsResourceCost_{rc}$.

Below are definitions and calculations for the variables used in Equation 4.11.4.

C—The number of trip types, 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/other, and misdemeanors.

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

B—The 50 years after treatment (the period over which we model the consequences of crime).

$CjsEvent_{bcf}$—Variable indicating if and when a criminal justice resource is used 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, $CrimeEvent_{ycf}$, to indicate when a resource is used. Each event matrix occurs within the recidivism follow-up period, *f*, for each trip type, *c*, and within the 50 years following treatment *b*. For criminal justice system events that occur over multiple years (e.g., prison), we incorporate length of stay information from Exhibit 4.11.4 into the event matrix.

$CjsResourcePr_{rc}$.—The probability that a criminal justice resource, *r*, will be used for a specific trip type, *c*. See Exhibit 4.11.4. For example, not all offenders who are convicted of a crime will necessarily receive a prison sentence.

$CjsResourceCost_{rc}$—The per unit marginal costs of each criminal justice resource as estimated in Section 4.11d of this Chapter and as shown in Exhibits 4.11.9 and 4.11.27.

$CjsResourceCostEsc_{r}$—The calculated real escalation rate of the unit marginal costs of each criminal justice resource as shown in Exhibit 4.11.9.

TotalTrips—The average number of trips through the criminal justice system during the follow-up period for each population.

TripTiming$_f$—Among those who offend during the follow-up period *f*, the probability that a trip happens in year *f*. The sum of *TripTiming*$_f$ equals 1.0.

TripTypePr$_c$— Among those who are convicted, the probability that at least one of the *TotalTrips* is of trip type is *c*. See Exhibit 4.11.1.

UnitΔ_f—The change in the probability of being convicted for a crime versus not being convicted in year *f*. This number is calculated using our effect size methods applied to 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 4.11.1. Different recidivism base rates are used depending on the specific population that receives a given program.

Victimizations Avoided. Using information from Exhibits 4.11.4, 4.11.8, and 4.11.29, we estimate the number of victimizations avoided and victimization costs avoided using the following equation:

$$(4.11.5)\quad Victim\$_b = \sum_{c=1}^{C}\sum_{f=1}^{F}\left[CjsEvent_{bcf} \times VictimVolume_c \times VictimCost_{rc} \times TotalTrips \times TripTiming_f \times TripTypePr_c \times Unit\Delta_f\right]$$

Below are definitions and calculations for the variables used in Equation 4.11.5 unless otherwise defined in the aforementioned section.

VictimVolume$_c$— Victimizations are shown in Exhibit 4.11.29.

$$(4.11.6)\quad VictimVolume_c = \sum_{v=1}^{V}(observed\ victims_v + unobserved\ victims_v)$$

VictimCost$_c$ —The per-unit cost of crime to victims as estimated in Section 4.11 of this Chapter and as shown in Exhibit 4.11.29.

Total Crime Costs. Using Equations 4.11.4 and 4.11.5 we discount the sum of the change in resources and victimization costs across different types of trips and time using the following equation:

$$(4.11.7)\quad Crime = \sum_{b=tage}^{B+tage}\sum_{r=1}^{10}\frac{\left(CjsResource\$_{ry} + Victim\$_y\right)}{(1+dis)^{(b-tage+1)}}$$

4.11g 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 the expected effect size and the estimated error 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.

4.11h 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 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 is described in a separate WSIPP report.[229]

There is a body of research literature on the effect of incarceration rates on crime.[230] Many of the studies addressing this relationship in the U.S. construct models using state-level data over a number of years to estimate the parameters of an equation of this general form:

$$(4.11.8) \quad C_{tsy} = a + b(ADP_{sy}) + c(X_{sy}) + e$$

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.[231] Many of these studies use data at the city or county level to estimate the parameters of an equation, such as the following:

$$(4.11.9) \quad C_{tcy} = a + b(POL_{cy}) + c(X_{cy}) + e$$

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 4.11.8 and 4.11.9 likely have unit roots, especially with state-level data.[232] Thus, to help avoid estimating spurious relationships, some authors estimate Equations 4.11.8 and 4.11.9 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 4.11.8 and 4.11.9 and in omitted variables bias.[233] 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 on 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 4.11.30 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 4.11.32 displays the meta-analytic results for prison length of stay on criminal recidivism.

[229] 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.
[230] 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.
[231] 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.
[232] See, for example, Marvell, (2010). See also, Spelman, W. (2008). Specifying the relationship between crime and prisons. *Journal of Quantitative Criminology, 24*, 149-178.
[233] Durlauf, S.N., & Nagin, D.S. (2010). *The deterrent effect of imprisonment NBER 5/07/10.*

Exhibit 4.11.30

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

Note:
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 4.11.30 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 4.11.10) and the effect of police on crime (Equation 4.11.11) using the following equations:

$$(4.11.10)\quad \Delta C_t = \frac{E_t \times \left(\frac{C_t}{ADP}\right)}{RR_t} \qquad (4.11.11)\quad \Delta C_t = \frac{E_t \times \left(\frac{C_t}{POL}\right)}{RR_t}$$

In Equations 4.11.10 and 4.11.11, 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 4.11.30; 2) the reported level of crime, *C*, for a particular crime type, *t*, as reported in Exhibit 4.11.31; 3) the incarceration rate, *ADP* (18,057), or the level of police employment, *POL* (10,502); and 4) the reporting rate to police by crime victims, *RR*, for a particular type of crime, *t*, as calculated from in Exhibit 4.11.6. 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 in Section 4.11e, the UCR definition of certain 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 4.11.6), to more closely align the UCR definitions with current law definitions in Washington, using the following equation:

$$(4.11.12)\quad C_t = UCR_t \times UCRAdj_t$$

In this analysis, we implement Equations 4.11.10 and 4.11.11 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 4.11.30. Therefore, we modify Equations 4.11.10 and 4.11.11 as follows:

$$(4.11.13)\quad \Delta C_v = \frac{(E_v \times R_v \times P_v) \times \left(\frac{C_v}{ADP}\right)}{RR_v} \qquad (4.11.14)\quad \Delta C_v = \frac{(E_v \times R_v \times P_v) \times \left(\frac{C_v}{POL}\right)}{RR_v}$$

$$(4.11.15)\quad \Delta C_p = \frac{(E_p \times R_p \times P_p) \times \left(\frac{C_p}{ADP}\right)}{RR_p} \qquad (4.11.16)\quad \Delta C_p = \frac{(E_p \times R_p \times P_p) \times \left(\frac{C_p}{POL}\right)}{RR_p}$$

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.[234] 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 4.11.13-4.11.16. 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 4.11.30 as the mean value and examine cohort-to-cohort variation to set the minimum and maximum values.

Exhibit 4.11.31
Three–Year Recidivism Rates of Offenders Released from Prison in Washington State,
Fiscal Years 2002 to 2004

Risk for re-offense category	Number of offenders	Recidivism for a violent felony offense		Recidivism for a property felony offense	
		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 4.11.30, other offenses, such as drug offenses, are not included in this definition.

The Policy Adjustment, *P*. Equations 4.11.13, 4.11.14, 4.11.15, and 4.11.16 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: 1) the probability of going to prison given a conviction and 2) 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 4.11.13 and 4.11.14 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-

[234] Barnoski & Drake (2007).

crime elasticity measured with the studies we include in the meta-analytic results displayed in Exhibit 4.11.31.[235] 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 4.11.32.

Exhibit 4.11.32

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

Note:
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 4.11.33. 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 4.11.32. 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 provide 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 U.S., prison length of stay increased by about four months, or about 17%, according to the U.S. 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 4.11.33. This produces an elasticity of -0.202. Since the elasticity for total UCR crime from our meta-analysis reported in Exhibit 4.11.30 is -0.26, a simple policy multiplier to use to analyze length of stay policy changes with Equations 4.11.13 and 4.11.14 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 4.11.30 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 are risk and uncertainty around this estimate, in the Monte Carlo simulation we model a triangular probability density distribution with lower and higher values in addition to the modal value of 0.776.

[235] 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 4.11.33

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, U.S., 1986 to 2009[1]	+4
(2)	Percentage 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)	Percentage change in recidivism rates[6]	-3.36%
(8)	Elasticity: percentage change in recidivism rate per percentage 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, U.S. 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 4.11.32.

[3] We assume a linear effect size and multiply the effect size from step (3), multiplied by 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 4.11.30, 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.[236] 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 4.11.14 and 4.11.16 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 4.11.34 and follow this computational process:

$$(4.11.17) \quad PM_t = \frac{ME_t + \frac{(HSES_t \times SD_t \times \overline{POP})}{\overline{\overline{POL}}}}{ME_t}$$

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).[237] 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

[236] Nagin (2013).

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

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.

Exhibit 4.11.34

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]		29,852

Notes:

[1] Marginal effect (E*C/POL) calculated with an elasticity, E, multiplied by 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 4.11.30.

[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), multiplied by the standard deviation in crime rates for Washington jurisdictions.

[5] The factor in footnote 4, multiplied by 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 4.11.13, 4.11.14, 4.11.15, and 4.11.16 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 4.11.18, 4.11.19, 4.11.20, and 4.11.21 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 policy 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, multiplied by R, multiplied by 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.

$$(4.11.18)\ \Delta C_v = \frac{\sum_{a=1}^{|\Delta ADP_T|} (E_v \times R_v \times P_v) \times \frac{[C_{v(a)} + (\Delta C_{v(a-1)})]}{(ADP_T \pm a)}}{RR_v}$$

$$(4.11.19)\ \Delta C_v = \frac{\sum_{a=1}^{|\Delta POL_T|} (E_v \times R_v \times P_v) \times \frac{[C_{v(a)} + (\Delta C_{v(a-1)})]}{(POL_T \pm a)}}{RR_v}$$

$$(4.11.20)\ \Delta C_p = \frac{\sum_{a=1}^{|\Delta ADP_T|} (E_p \times R_p \times P_p) \times \frac{[C_{p(a)} + (\Delta C_{p(a-1)})]}{(ADP_T \pm a)}}{RR_p}$$

$$(4.11.21)\ \Delta C_p = \frac{\sum_{a=1}^{|\Delta POL_T|} (E_p \times R_p \times P_p) \times \frac{[C_{p(a)} + (\Delta C_{p(a-1)})]}{(POL_T \pm a)}}{RR_p}$$

For example, for a policy that decreases prison ADP by 100 units, Equations 4.11.18 and 4.11.20 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 4.11.18, 4.11.19, 4.11.20, and 4.11.21. 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:

$$(4.11.22)\ \Delta RC_v = \Delta C_v \times RR_v$$

$$(4.11.23)\ \Delta RC_p = \Delta C_p \times RR_p$$

Exhibit 4.11.35
Washington Criminal Justice System Resources

Washington court and criminal justice numbers	Murder	Felony sex crimes	Robbery	Aggravated assault	Felony property	Years
Number of arrests, adult and juvenile	156	1,409	2,129	6,134	41,165	2011-2014
Number of trips, adult and juvenile	220	1,065	1,020	6,496	9,632	2011-2015
Number of convictions, adult and juvenile	264	1,747	1,335	8,651	17,995	2011-2015

Number of Arrests, Adult and Juvenile. Adult and juvenile felony conviction data are obtained from FBI UCR Crime publications.[238]

Number of Trips, Adult and Juvenile. Adult and juvenile felony conviction trips are calculated using the WSIPP Criminal Justice System Database.

Number of Counts, Adult and Juvenile. Adult and juvenile felony convictions are calculated using the WSIPP Criminal Justice System Database.

[238] Information for Washington taken from *Crime in the United States* Data Series FBI Table 69.

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

$$(4.11.24)\ \Delta Victim\$ = \Delta C_v \times VictimPerUnit\$_v + \Delta C_p \times VictimPerUnit\$_p$$

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 4.11.11, ΔC_v and ΔC_p multiplied by, 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 is 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 is calculated using the following equations:

$$(4.11.25)\ \Delta Police\$ = \Delta RC_v \times \frac{Arrest_v}{RC_v} \times PolicePerArrest\$_v + \Delta RC_p \times \frac{Arrest_p}{RC_p} \times PolicePerArrest\$_p$$

$$(4.11.26)\ \Delta Court\$ = \Delta RC_v \times \frac{Trip_v}{RC_v} \times CourtPerTrip\$_v + \Delta RC_p \times \frac{Trip_p}{RC_p} \times CourtPerTrip\$_p$$

$$(4.11.27)\ \Delta Jail\$ = \Delta RC_v \times \frac{JailLOS_v}{RC_v} \times JailPerYear\$_v + \Delta RC_p \times \frac{JailLOS_p}{RC_p} \times JailPerYear\$_p$$

$$(4.11.28)\ \Delta Prison\$ = \Delta RC_v \times \frac{PrisonLOS_v}{RC_v} \times PrisonPerYear\$_v + \Delta RC_p \times \frac{PrisonLOS_p}{RC_p} \times PrisonPerYear\$_p$$

$$(4.11.29)\ \Delta JailCS\$ = \Delta RC_v \times \frac{JailCSLOS_v}{RC_v} \times JailCSPerYear\$_v + \Delta RC_p \times \frac{JailSuperLOS_p}{RC_p} \times JailCSPerYear\$_p$$

$$(4.11.30)\ \Delta PrisonCS\$ = \Delta RC_v \times \frac{PrisonCSLOS_v}{RC_v} \times PrisonCSPerYear\$_v + \Delta RC_p \times \frac{PrisonCSLOS_p}{RC_p} \times PrisonCSPerYear\$_p$$

For each segment of the criminal justice system, the change in expenses is the sum of the change in the number of reported violent and property victimizations from Equations 4.11.22 and 4.11.23, ΔRC_v and ΔRC_p multiplied by, respectively, the probability that a reported crime uses resources in each criminal justice segment, multiplied by the marginal cost of that segment per violent and property victimization. For jail and prison length of stay and the length of stay on community supervision for jail-based and post-prison-based segments, the parameters are conditional on the probability of a trip 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.

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 our 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.aa.7 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 WSIPP model. The comparison was made in the January 2016 edition of the model and values may not reflect current model estimates.

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.

Exhibit 4.12.1

WSIPP's Benefit-Cost Values Compared to Other Researchers

Outcome	Comparison study			WSIPP benefit-cost model		Common year of dollars and discount rate
	Study	Key result	Notes on study	WSIPP result	Comparison note	
Cigarette smoking, total lifetime costs	Sloan, F.A., Ostermann, J., Picone, G. Conover, C., & Taylor, D.H. Jr. (2004). *The price of smoking.* MIT Press.	$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. (2015). Health care expenses in relation to obesity and smoking among U.S. 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 high school $386,392 if continue on to more education at rate of high school grads	Ages 20-67. 2004 dollars. Uses cross-sectional differences in CPS. GEDs treated as high school graduates, excludes prison population and 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

Outcome	Comparison study			WSIPP benefit-cost model		Common year of dollars and discount rate
	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 high school graduates 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. (2012). Economic burden of child maltreatment in the U.S. 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].	$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. (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

When a program has multiple outcomes *in one or more of the constructs* described in Section 5.2, we apply trumping rule 1. We then apply either trumping rule 2, 3, or 4:

- ✓ **Trumping Rule 1: Direct over Indirect.** In situations where there 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. This rule overrides the previous three rules.

  As noted in this document, the WSIPP benefit-cost model monetizes changes to outcomes measured in one of two ways: 1) directly from program evaluations that measure an outcome of interest or 2) from "linkage" studies that measure how a change in one outcome leads to a change in a second outcome (see expression in Section 2.1 and 3.3). For example, many program evaluations measure a program's impact on crime, which then can be directly valued—a change in the likelihood of crime leads to a change in the expected dollars for the criminal justice system and for victims. Alternatively, many program evaluations measure a program's impact on high school graduation. While we can directly value high school graduation via expected changes in labor market earnings, we can also *indirectly* measure that program's impact on crime via the "linkage" research literature, which allows us to approximate the magnitude of the causal impact of high school graduation on crime participation. The program's effect on high school graduation *indirectly* leads to an impact on crime, which can also be valued.

  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. Per the fourth trumping rule, we would only monetize the high school graduation outcome because it was directly measured.

- ✓ **Trumping Rule 2: 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. We then select the outcome with the largest present value of benefits and drop the other outcomes.[239] 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 while dropping the other outcome from the benefit-cost analysis.

[239] 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 Biggest Winner and the Biggest Loser.** When a topic has at least one favorable and at least one unfavorable (i.e., iatrogenic) alternative outcome, we determine the expected present value of benefits of each favorable alternative outcome and the expected present value of losses of each unfavorable outcome. 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 add them together. Any other competing outcomes are dropped from the benefit-cost analysis.
- ✓ **Trumping Rule 4: The Biggest 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 largest magnitude 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 largest magnitude present value loss, dropping the other outcome from the benefit-cost analysis.

5.2 Underlying Constructs

As noted, certain outcomes are likely to be alternative ways of measuring the same construct. In the WSIPP benefit-cost model, we have identified the following types of outcomes that are alternative ways of measuring the same construct. For a complete listing of all outcomes monetized in the WSIPP benefit-cost model and the underlying construct they reflect, see Exhibit 5.2.1.

- For outcomes associated with labor market earnings, we assume that the labor market gains from 1) increases in academic achievement; 2) increases in academic attainment; 3) decreases in substance abuse; 4) decreases in mental health conditions; 5) decreases in health conditions; and 6) reductions in child abuse and neglect reflect different measures of the same underlying construct that affects labor market performance.
- For outcomes that change the probability of mortality, we assume that changes in mortality from 1) decreases in substance abuse; 2) decreases in mental health conditions; 3) increases in infant health; 4) decreases in health conditions; 5) decreases in the likelihood of falling; and 6) reductions in child abuse and neglect all reflect different ways of approximating the same construct.
- For health care outcomes, we assume that health care costs stemming from changes in 1) high school graduation; 2) substance abuse; 3) mental health conditions; 4) health conditions; 5) falls; 6) birth outcomes; and 7) utilization of specific health care services all reflect different measures of the same underlying construct that affect health care costs.
- For outcomes that affect the amount of time spent in higher education, we assume that changes in costs from 1) the likelihood of graduating high school; 2) the likelihood of graduating from 2-year or 4-year college programs; and 3) persisting in higher education programs reflect different measures of the same construct that affects participation in higher education.
- For outcomes that affect property loss, we consider that lost property resulting from either alcohol use disorder or problem alcohol use both reflect the same construct.
- Finally, there are a number of outcomes that stand alone, i.e., the WSIPP model only has one outcome that measures each construct:
    - Crime
    - K-12 grade repetition
    - K-12 special education
    - Cash assistance
    - Food assistance

Exhibit 5.2.1 Construct and Outcome Relationship

Construct	Outcomes measuring that construct	
Criminal justice system	✓ Crime	
Child welfare system	✓ Child abuse and neglect	✓ Out-of-home placements
K12 system	✓ Special Education	✓ Grade retention
Human capital labor market earnings	✓ Earnings ✓ Employment ✓ Child abuse and neglect ✓ Student test scores ✓ High school graduation ✓ Higher education graduation ✓ Persistence in higher education ✓ Alcohol use disorder ✓ Problem alcohol use ✓ Cannabis use disorder	✓ Illicit drug use disorder ✓ Opioid use disorder ✓ Regular smoking ✓ Depression ✓ Anxiety ✓ Post-traumatic stress disorder ✓ Diabetes ✓ Obesity
Mortality	✓ Alcohol use disorder ✓ Problem alcohol use ✓ Illicit drug use disorder ✓ Opioid use disorder ✓ Regular smoking ✓ Depression	✓ Diabetes ✓ Obesity ✓ Falls ✓ Child abuse and neglect ✓ Infant mortality
Health care costs	✓ Utilization of specific services, represented by a sum of: o Emergency department visits, o Hospitalizations, o Psychiatric hospitalizations, and o Hospital readmissions ✓ High school graduation ✓ Alcohol use disorder ✓ Problem alcohol use ✓ Cannabis use disorder ✓ Illicit drug use disorder ✓ Opioid use disorder ✓ Regular smoking ✓ Attention deficit hyperactivity disorder ✓ Anxiety ✓ Disruptive behavior disorder	✓ Internalizing symptoms ✓ Externalizing symptoms ✓ Depression ✓ Posttraumatic stress disorder ✓ Diabetes ✓ Obesity ✓ Falls ✓ Cesarean sections ✓ Low birthweight births ✓ Very low birthweight births ✓ Neonatal intensive care unit use ✓ Preterm births ✓ Small for gestational age births
Higher education costs	✓ High school graduation ✓ Persistence in higher education	✓ Higher education graduation
Property loss	✓ Alcohol use disorder	✓ Problem alcohol use

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 6.1.1.

$$(6.1.1)\quad NPV_{tage} = \sum_{y=tage}^{N} \frac{Q_y \times P_y - C_y}{(1+Dis)^y}$$

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 the 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 similarly. 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 6.1.1.

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 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 Estimate 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 benefits exceed costs, 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)?

Since 2013, the Washington State 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;" in other words, benefits must exceed costs in at least 75% of the 10,000 Monte Carlo simulation runs. If benefits exceed costs 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 use the mean effect size and random-effects standard error to create a normal probability density distribution of possible "unit changes" caused by the program (described in Chapter 3).

Linked Effect Sizes. Chapters 2 and 3 also describe how the model uses estimates of how 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 random effects standard error to create a normal probability density distribution of possible "unit changes" caused by the program (described in Chapter 3).

Discount Rates. Three different rates of discount (low, modal, and high) are used to evaluate future benefits and costs in present value terms. In a single run of the model, the modal discount rate is used. In Monte Carlo simulation mode, the discount rate is sampled from a triangular probability density distribution. A discussion of the discount rate parameters can be found in Section 4.aab.

The mean or modal values for many other model inputs are varied in a Monte Carlo run and include the following:

- Crime victimization costs and Child abuse and neglect victimization costs—triangular distribution described in Section 4.11e
- Ratios of other victims per trip—triangular distribution—See Exhibit 4.11.6
- Criminal justice system costs—triangular distribution—See Exhibit 4.11.9
- Crime police and prison elasticities—normal distribution—See Exhibit 4.11.30
- Value of a statistical life—triangular distribution—See Exhibit 4.1.3
- Child abuse and neglect system costs—triangular distribution—See Exhibit 4.10.5
- Deadweight cost of taxation—triangular distribution—See Section 4.aae
- Labor market earnings ratios—drawn from distributions described in corresponding Exhibits:
  - Substance abuse/dependence—See Exhibit 4.5.9
  - Mental health disorders—See Exhibit 4.6.3
  - Child abuse and neglect—See Exhibit 4.10.8
  - Health conditions (obesity and diabetes) —See Exhibit 4.7.20
- Expected higher education cost escalation—triangular distribution—See Section 4.8c
- Expected health care cost escalation—triangular distribution—See Section 4.3a
- Expected health care costs:
  - Mental health disorders—normal distribution—See Exhibit 4.6.4
  - Substance use disorders—normal distribution—See Exhibit 4.5.10
  - Health Care utilization measures—normal distribution—See Exhibits 4.3.3 – 4.3.6
  - Falls health care costs—triangular distribution—See Exhibit 4.3.9
  - Health care disorder—normal distribution—See Exhibit 4.7.21
- Labor market earnings from one standard deviation increase in test scores—normal distribution—See Exhibit 4.8.1
- Causal links between educational attainment and earnings—normal distribution:
  - Between high school graduation and labor market earnings for varying education levels—See Exhibit 4.8.5
  - Between higher educational enrollment/graduation at a 2/4-year institution and labor market earnings—See Exhibit 4.8.4
  - Between years of persistence in a 2/4-year institution and labor market earnings—See Exhibit 4.8.13
- Human capital economic externalities of education—triangular distribution—See Exhibit 4.8.1
- Expected system costs of Child Abuse and Neglect—triangular distribution—See Exhibit 4.10.5

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:

$$(7.2.1) \quad RandomDraw = RAND()$$

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:

$$(7.2.2) \quad EffectSize_r = NORMINV(RandomDraw, EffectSizeMean, EffectSizeStandardError)$$

Other types of probability distributions are computed similarly.

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.

$$(7.2.3)\quad If\ RandomDraw < \frac{(Mode - Min)}{(Max - Min)}\ then\ DISRATE = Min + \sqrt{RandomDraw \times (Mode - Min) \times (Max - Min)}$$

$$(7.2.4)\quad If\ RandomDraw \geq \frac{(Mode - Min)}{(Max - Min)}\ then\ DISRATE = Max - \sqrt{(1 - RandomDraw) \times (Max - Mode) \times (Max - Min)}$$

Chapter 8: The WSIPP Portfolio Tool

WSIPP constructed an analytical portfolio tool for the Washington State 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[240] but has been expanded to include several new outcomes, not just those relevant to criminal justice.[241] 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, 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 Sections 4.2, 4.11, 4.10, and 4.8.

Key parameters that are allowed to vary in the individual benefit-cost model are described in Section 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.

[240] 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.

[241] 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.

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: 1) the current annual funded participants in each program, 2) the maximum number of annual eligible participants, and 3) how effective the 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 a 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 that 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 than 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 ($DRFactorpy$).

$$(8.3.1) \quad \Delta Person_i = \sum_{p=1}^{N} \sum_{y=1}^{i} \Delta Person_{p(i-y+1)} * AddSlots_{py} * AdjFactor_p * DRFactor_{py}$$

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.

$$(8.3.2) \quad DRFactor_y = \frac{FV\left(RATE(99,0,-1,DR),\left(\frac{(Current + AddSlots_y)}{MaxCap} * 100,0,-1\right)\right) + FV\left(RATE(99,0,-1,DR),\left(\frac{Current}{MaxCap} * 100,0,-1\right)\right)}{2}$$

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-averse 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 8.4.1

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

Note:

* 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 Appendices

Appendices

AI. 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 A.I.1 displays the meta-analytic results of each linkage we have estimated; Exhibit A.I.2 shows the individual studies for each linkage; and Exhibit A.I.3 is a list of citations for all of the studies in these meta-analyses of linked outcomes.

Exhibit A.I.1

Linked Outcomes Meta-Analytic Estimates of Standardized Mean Difference Effect Sizes

Estimated causal links between outcomes	No. of effect sizes	Meta-analytic results before adjusting effect sizes								Adjusted effect size and standard error used in the benefit-cost analysis		Age of link measurement	Age at which relationship begins
		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…													
Crime	See externalizing composite												
Grade retention	4	0.466	0.048	0.000	3.021	0.388	0.466	0.049	0.000	**0.466**	**0.049**	**8**	**17**
High school graduation	4	-0.311	0.025	0.000	0.108	0.991	-0.311	0.025	0.000	**-0.311**	**0.025**	**18**	**18**
Special education	See externalizing composite												
Test scores-academic	See externalizing composite												
Alcohol disorder, leading to…													
Crime	2	0.253	0.052	0.000	1.093	0.296	0.249	0.057	0.000	**0.249**	**0.057**	**30**	**1**
Alcohol (problem use), leading to…													
Crime	3	0.260	0.046	0.000	0.306	0.858	0.260	0.046	0.000	**0.260**	**0.046**	**30**	**1**
High school graduation	7	-0.112	0.035	0.002	12.383	0.054	-0.166	0.061	0.007	**-0.166**	**0.061**	**18**	**18**
Alcohol use < 14 years of age, leading to…													
Crime	4	0.133	0.034	0.000	0.306	0.959	0.133	0.034	0.000	**0.133**	**0.034**	**20**	**13**
High school graduation	6	-0.034	0.017	0.044	8.406	0.135	-0.039	0.030	0.201	**-0.039**	**0.030**	**18**	**18**
Alcohol use < 18 years of age, leading to…													
Crime	See alcohol use < 14 years of age (note: alcohol use < 18 has unique age references)											**20**	**15**
High school graduation	See alcohol use <14 years of age												
Anxiety, leading to…													
Grade retention	See internalizing composite												
High school graduation	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**	**16**	**17**
High school graduation	3	-0.213	0.068	0.002	0.841	0.657	-0.213	0.068	0.002	**-0.213**	**0.068**	**18**	**18**
Tobacco (regular use)	1	0.052	0.137	0.706	0.000	0.000	0.052	0.137	0.706	**0.052**	**0.137**	**20**	**1**
Births to < 18 mother (mother effect), leading to…													
High school graduation	4	-0.109	0.066	0.097	1.865	0.601	-0.109	0.066	0.097	**-0.109**	**0.066**	**18**	**18**
Public Assistance	2	0.107	0.101	0.287	0.047	0.828	0.107	0.101	0.287	**0.107**	**0.101**	**25**	**18**
Cannabis use < 14 years of age, leading to…													

Estimated causal links between outcomes	No. of effect sizes	Meta-analytic results before adjusting effect sizes								Adjusted effect size and standard error used in the benefit-cost analysis		Age of link measurement	Age at which relationship begins
		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		
Crime	1	0.271	0.130	0.038	0.000	0.000	0.271	0.130	0.038	**0.271**	**0.130**	**20**	**13**
High school graduation	13	-0.180	0.016	0.000	109.830	0.000	-0.235	0.064	0.000	**-0.235**	**0.064**	**18**	**18**
Cannabis use < 18 years of age, leading to...													
Crime	See cannabis use < 14 years of age (note: cannabis use < 18 has unique age references)											**20**	**15**
High school graduation	See cannabis use < 14 years of age												
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**	**25**	**1**
Child abuse & neglect, leading to...													
Alcohol (disordered use)	6	0.171	0.028	0.000	7.590	0.180	0.172	0.046	0.000	**0.172**	**0.046**	**25**	**18**
Anxiety (incl. OCD)	3	0.298	0.052	0.000	17.366	0.000	0.325	0.166	0.051	**0.325**	**0.166**	**20**	**18**
Crime	11	0.532	0.034	0.000	35.330	0.000	0.542	0.071	0.000	**0.542**	**0.071**	**20**	**18**
Depression	8	0.305	0.028	0.000	22.675	0.002	0.293	0.058	0.000	**0.293**	**0.058**	**20**	**18**
Disruptive behavior	1	0.460	0.391	0.239	0.000	0.000	0.460	0.391	0.239	**0.460**	**0.391**	**12**	**12**
Grade retention	1	0.446	0.102	0.000	0.000	0.000	0.446	0.102	0.000	**0.446**	**0.102**	**12**	**17**
High school graduation	5	-0.412	0.048	0.000	14.308	0.006	-0.405	0.098	0.000	**-0.405**	**0.098**	**18**	**18**
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**	**21**	**18**
Obesity	5	0.022	0.018	0.242	9.052	0.060	0.042	0.039	0.283	**0.042**	**0.039**	**35**	**18**
PTSD	1	0.836	0.199	0.000	0.000	0.000	0.836	0.199	0.000	**0.836**	**0.199**	**18**	**18**
Special education	1	0.389	0.036	0.000	0.000	0.000	0.389	0.036	0.000	**0.389**	**0.036**	**8**	**5**
Test scores-academic	2	-0.270	0.062	0.000	2.278	0.320	-0.268	0.067	0.000	**-0.268**	**0.067**	**17**	**17**
Tobacco (regular use)	1	0.387	0.123	0.002	0.000	0.000	0.387	0.123	0.002	**0.387**	**0.123**	**20**	**18**
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**	**18**	**18**
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**	**18**	**18**
Depression, leading to...													
Grade retention	See internalizing composite												
High school graduation	See internalizing composite												
Diabetes, leading to...													
Nursing home	8	0.212	0.008	0.000	20.497	0.005	0.210	0.046	0.000	**0.210**	**0.046**	**75**	**1**
Disruptive behavior, leading to...													

Estimated causal links between outcomes	No. of effect sizes	Meta-analytic results before adjusting effect sizes								Adjusted effect size and standard error used in the benefit-cost analysis		Age of link measurement	Age at which relationship begins
		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		
Crime	See externalizing composite												
Grade retention	4	0.273	0.055	0.000	1.155	0.764	0.273	0.055	0.000	**0.273**	**0.055**	**16**	**17**
High school graduation	6	-0.4210	0.0260	0.0000	6.4280	0.2670	-0.4317	0.0339	0.0000	-0.4317	0.0339	18	18
Special education	See externalizing composite												
Test scores-academic	See externalizing composite												
Externalizing composite (includes conduct disorder & ADHD), leading to...													
Crime	8	0.328	0.035	0.000	12.107	0.097	0.340	0.056	0.000	**0.340**	**0.056**	**20**	**1**
High school graduation	3	-0.225	0.029	0.000	1.261	0.532	-0.225	0.029	0.000	**-0.225**	**0.029**	**18**	**18**
Special education	2	0.398	0.091	0.000	0.047	0.828	0.398	0.091	0.000	**0.398**	**0.091**	**16**	**1**
Test scores-academic	5	-0.145	0.020	0.000	35.066	0.000	-0.185	0.076	0.015	**-0.185**	**0.076**	**13**	**1**
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**	**25**	**18**
Illicit drugs, leading to													
Crime	2	0.3043	0.0559	0.0000	0.0722	0.7881	0.3043	0.0559	0.0000	**0.3043**	**0.0559**	**30**	**1**
Internalizing composite (includes depression & anxiety), leading to...													
Grade retention	2	0.266	0.052	0.000	0.564	0.453	0.266	0.052	0.000	**0.266**	**0.052**	**16**	**17**
High school graduation	7	-0.109	0.027	0.000	9.142	0.166	-0.117	0.037	0.002	**-0.117**	**0.037**	**18**	**18**
Low birth weight (< 2,500 g), leading to...													
Infant mortality	1	1.437	0.078	0.000	0.000	0.000	1.437	0.078	0.000	**1.437**	**0.078**	**1**	**1**
Obesity, leading to...													
Nursing home	3	0.177	0.030	0.000	0.840	0.657	0.177	0.030	0.000	**0.177**	**0.030**	**75**	**1**
Opioids, leading to													
Crime	See illicit drugs											**30**	**1**
Preterm birth (< 37 weeks gestation), leading to...													
Infant mortality	1	1.1034	0.0719	0.0000	0.0000	0.0000	1.1034	0.0719	0.0000	**1.1034**	**0.0719**	**1**	**1**
Small for gestational age, leading to...													
Infant mortality	1	0.7944	0.0777	0.0000	0.0000	0.0000	0.7944	0.0777	0.0000	**0.7944**	**0.0777**	**1**	**1**
Smoking regularly <14 years of age, leading to...													
High school graduation	5	-0.394	0.016	0.000	14.536	0.006	-0.351	0.055	0.000	**-0.351**	**0.050**	**18**	**18**
Smoking regularly <18 years of age, leading to...													
High school graduation	See smoking regularly < 14 years of age												
Very low birthweight (< 1,500g), leading to...													
Infant mortality	1	2.020	0.132	0.000	0.000	0.000	2.020	0.132	0.000	**2.020**	**0.132**	**1**	**1**

Exhibit A.I.2

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Composite record Id	Unadjusted effect size	No. in test condition group	No. in control group	Inverse variance weight - fixed effects	Inverse variance weight· random effects	WSIPP multipliers	Adjusted effect size
ADHD		**Crime**							
See Externalizing composite									
ADHD		**Grade retention**							
605	Fletcher & Wolfe, 2008		0.453	261	2,643	53.979	53.773	1.000	**0.453**
12756	Galera et al., 2009		0.597	163	1,101	92.785	92.177	1.000	**0.597**
13157	Currie & Stabile, 2007		0.347	359	3,232	99.276	98.581	1.000	**0.347**
13170	Currie & Stabile, 2007		0.468	582	5,240	179.410	177.153	1.000	**0.468**
ADHD		**High school graduation**							
610	Fletcher & Wolfe, 2008		-0.305	262	2,645	93.840	93.840	1.000	**-0.305**
9848	Breslau et al., 2008		-0.322	486	5,100	201.047	201.047	1.000	**-0.322**
12755	Galera et al., 2009		-0.369	71	643	22.603	22.603	1.000	**-0.369**
13712	Breslau et al., 2011		-0.309	2,966	26,696	1328.593	1328.593	1.000	**-0.309**
ADHD		**Special education**							
See Externalizing composite									
ADHD		**Test scores-academic**							
See Externalizing composite									
Alcohol Disorder		**Crime**							
17705	Popovici et al., 2012		0.280	756	8,820	290.072	236.568	1.000	**0.280**
13742	Van Dorn et al., 2012		0.145	2,946	31,707	74.745	70.629	1.000	**0.145**
Alcohol (problem use)		**Crime**							
17706	Popovici et al., 2012		0.251	756	8,820	282.446	282.446	1.000	**0.251**
18554	Hill et al., 2000		0.156	242	566	16.301	16.301	1.000	**0.156**
18589	Viner & RM, 2007		0.285	874	4,037	173.225	173.225	1.000	**0.285**
Alcohol (problem use)		**High school graduation**							
10570	Renna, 2008		-0.033	654	1,310	247.904	64.671	1.000	**-0.033**
11955	Chatterji et al., 2005		-0.349	105	1,002	47.508	30.790	1.000	**-0.349**
7910	Dee & Evans, 2003		-0.222	1,717	5,749	70.274	38.973	1.000	**-0.222**
18258	Hawkins et al., 2013		-0.028	86	5,314	25.002	19.446	1.000	**-0.028**
18383	Yan & Brocksen, 2013		-0.303	290	1,523	52.728	32.901	1.000	**-0.303**
18461	Chatterji, 2006		-0.070	1,695	5,909	344.501	69.775	1.000	**-0.070**
18462	Hill et al., 2000		-0.799	242	566	9.798	8.812	1.000	**-0.799**
Alcohol use < 14 years of age		**Crime**							
18527	Ellickson et al., 2003		0.131	2,523	846	632.541	632.541	1.000	**0.131**
18549	Green et al., 2011	18552	0.332	186	516	78.124	78.124	1.000	**0.332**
18550	Green et al., 2011	18552	0.095	186	516	70.528	70.528	1.000	**0.095**
18551	Green et al., 2011	18552	0.000	186	516	51.823	51.823	1.000	**0.000**
18588	Newcomb & McGee, 1989		0.235	549	298	19.977	19.977	1.000	**0.235**
18630	Wells et al., 2004	18633	0.166	729	224	170.926	170.926	1.000	**0.166**
18631	Wells et al., 2004	18633	0.166	729	224	170.926	170.926	1.000	**0.166**
Alcohol use < 14 years of age		**High school graduation**							
10569	Renna, 2008		-0.113	1,082	882	280.695	181.311	1.000	**-0.113**
7081	Ellickson et al., 1998		0.057	3,279	1,111	260.055	172.469	1.000	**0.057**
7088	Bray et al., 2000		0.177	1,144	248	36.593	34.152	1.000	**0.177**
7911	Dee & Evans, 2003		-0.250	1,419	4,330	57.952	52.060	1.000	**-0.250**
18460	Chatterji, 2006		-0.052	3,029	4,575	462.403	242.988	1.000	**-0.052**
13719	Breslau et al., 2011		-0.030	8,543	21,119	2380.849	421.438	1.000	**-0.030**
Alcohol use < 18 years of age		**Crime**							
See Alcohol Use <14 years of age									
Alcohol use < 18 years of age		**High school graduation**							
See Alcohol Use <14 years of age									
Anxiety		**Grade retention**							
See Internalizing composite									
Anxiety		**High school graduation**							
See Internalizing composite									
Births to < 18 Mother (child effect)		**Grade retention**							
7038	Angrist & Lavy, 1996		0.213	557	17,238	539.161	4.348	1.000	**0.213**
26598	Moore et al., 1997		0.245	77	199	24.214	3.711	1.000	**0.245**
12796	Levine et al., 2007	24348	-0.061	701	551	120.484	4.229	0.000	**0.000**
12797	Levine et al., 2007	24348	1.445	140	219	41.697	3.966	0.000	**0.000**
Births to < 18 mother (child effect)		**High school graduation**							
26612	Francesconi, 2008		-0.314	85	1,098	53.592	53.592	1.000	**-0.314**
12783	Hoffman & Scher, 2008		-0.205	644	337	86.842	86.842	1.000	**-0.205**
12794	Manlove et al., 2008		-0.150	221	461	73.754	73.754	1.000	**-0.150**
Births to < 18 mother (child effect)		**Tobacco (regular use)**							
12801	Francesconi, 2008		0.052	85	1,098	53.174	0.000	1.000	**0.052**
Births to < 18 mother (mother effect)		**High school graduation**							
11799	Fletcher & Wolfe, 2009		-0.241	563	148	71.146	71.146	1.000	**-0.241**
26715	Ashcraft et al., 2013	27225	-0.025	1,313	186	109.811	109.811	1.000	**-0.025**
27224	Ashcraft et al., 2013	27225	-0.049	1,313	186	109.738	109.738	1.000	**-0.049**
12785	Webbink et al., 2009		-0.065	77	77	25.412	25.412	1.000	**-0.065**
12800	Hoffman, 2008		-0.096	453	41	25.279	25.279	1.000	**-0.096**

Record Id	Citation	Composite record Id	Unadjusted effect size	No. in test condition group	No. in control group	Inverse variance weight - fixed effects	Inverse variance weight- random effects	WSIPP multipliers	Adjusted effect size
	Births to < 18 mother (mother effect)	**Public assistance**							
11800	Fletcher & Wolfe, 2009		0.137	564	149	35.007	35.007	1.000	**0.137**
12799	Hoffman, 2008		0.091	762	69	63.251	63.251	1.000	**0.091**
	Cannabis < 14 years of age	**Crime**							
18155	Green et al., 2010		0.271	185	517	58.900	NaN	1.000	**0.271**
	Cannabis < 14 years of age	**High school graduation**							
7079	Ellickson et al., 1998		-0.074	860	3,530	197.850	23.911	1.000	**-0.074**
7086	Bray et al., 2000		-0.508	677	715	26.314	13.375	1.000	**-0.508**
7151	Fergusson & Horwood, 1997		-0.385	180	755	73.897	19.881	1.000	**-0.385**
7159	Brook et al., 2002		-0.217	100	1,048	27.213	13.603	1.000	**-0.217**
7389	McCaffrey et al., 2009		-0.112	276	2,482	20.391	11.654	1.000	**-0.112**
18151	Green et al., 2010		-0.595	185	517	31.913	14.684	1.000	**-0.595**
18256	Hawkins et al., 2013		0.104	124	5,276	34.253	15.161	1.000	**0.104**
18320	Legleye et al., 2010		0.078	13,026	16,367	765.410	26.265	1.000	**0.078**
12749	Yamada et al., 1996		-0.179	75	597	12.922	8.760	1.000	**-0.179**
12804	van Ours & Williams, 2009		-0.198	5,931	5,862	1992.438	26.832	1.000	**-0.198**
12811	Horwood et al., 2010		-0.480	1,418	2,176	337.609	25.171	1.000	**-0.480**
12815	Horwood et al., 2010		-0.162	407	1,036	106.559	21.668	1.000	**-0.162**
12817	Horwood et al., 2010		-0.387	994	2,176	267.063	24.685	1.000	**-0.387**
	Cannabis use <18 years of age	**High school graduation**							
	See Cannabis <14 years of age								
	Cannabis use <18 years of age	**Crime**							
	See Cannabis <14 years of age								
	Cesarean Section	**Hospital readmissions**							
26749	Liu et al., 2002		0.379	483,263	2,169,463	10494.907	10494.907	1.000	**0.379**
	Child Abuse & Neglect	**Alcohol (disordered use)**							
10552	Scott et al., 2010		0.332	221	1,923	55.115	44.866	1.000	**0.332**
12408	Thornberry et al., 2010		0.171	170	645	134.214	86.240	1.000	**0.171**
12421	Shin et al., 2009		0.173	6,729	6,019	851.913	188.020	1.000	**0.173**
6768	Fergusson & Lynskey, 1997		0.409	118	111	23.929	21.770	1.000	**0.409**
22106	Horwitz et al., 2001		-0.058	315	271	88.989	65.011	1.000	**-0.058**
22107	Horwitz et al., 2001		0.214	322	239	85.847	63.318	1.000	**0.214**
	Child Abuse & Neglect	**Anxiety (incl. OCD)**							
21980	Springer et al., 2007		0.165	234	1,817	207.018	12.916	1.000	**0.165**
22232	Scott et al., 2010		0.649	221	1,923	101.776	12.134	1.000	**0.649**
12947	Fergusson et al., 2008		0.157	162	839	57.227	11.103	1.000	**0.157**
	Child Abuse & Neglect	**Crime**							
12406	Thornberry et al., 2010		0.342	170	645	82.553	23.587	1.000	**0.342**
6718	English et al., 2002		0.600	877	877	235.640	28.963	1.000	**0.600**
6749	Stouthamer-Loeber et al., 2001		0.379	52	104	23.086	13.587	1.000	**0.379**
6820	Lansford et al., 2007	22109	0.718	69	505	14.431	10.043	1.000	**0.718**
6821	Lansford et al., 2007	22109	0.634	69	505	25.082	14.255	1.000	**0.634**
6858	Mersky & Reynolds, 2007	6860	0.528	129	1,275	52.741	20.307	1.000	**0.528**
6859	Mersky & Reynolds, 2007	6860	0.364	129	1,275	46.281	19.272	1.000	**0.364**
7388	Lemmon, 1999		1.083	267	365	79.823	23.359	1.000	**1.083**
9057	Stouthamer-Loeber et al., 2002		0.635	83	179	31.622	16.154	1.000	**0.635**
21923	Cohen et al., 2004		0.530	51	579	30.475	15.849	1.000	**0.530**
21998	Currie & Tekin, 2012		0.414	512	1,704	147.389	26.978	1.000	**0.414**
22181	Kazemian, et al., 2011		0.477	202	50	14.509	10.080	1.000	**0.477**
22265	Allwood & Widom, 2013		0.389	676	520	167.288	27.578	1.000	**0.389**
	Child Abuse & Neglect	**Depression**							
10544	Scott et al., 2010		0.525	221	1,923	68.961	35.787	1.000	**0.525**
12409	Thornberry et al., 2010		0.158	170	645	134.264	47.870	1.000	**0.158**
12426	Fletcher, 2009	22116	0.211	196	3,801	80.425	38.646	1.000	**0.211**
12427	Fletcher, 2009	22116	0.382	168	3,880	80.355	38.630	1.000	**0.382**
12433	Springer et al., 2007		0.156	234	1,817	207.047	54.729	1.000	**0.156**
6811	Chapman et al., 2004	21916	0.437	2,850	6,610	587.147	66.028	1.000	**0.437**
6812	Chapman et al., 2004	21916	0.377	1,896	7,564	453.466	63.909	1.000	**0.377**
6850	Widom et al., 2007		0.145	676	520	139.298	48.495	1.000	**0.145**
22260	Brown et al., 1999		0.665	81	558	18.774	14.991	1.000	**0.665**
12946	Fergusson et al., 2008		0.266	162	839	76.537	37.725	1.000	**0.266**
	Child Abuse & Neglect	**Disruptive Behavior**							
6769	Fergusson & Lynskey, 1997		0.460	118	111	6.553	0.000	1.000	**0.460**
	Child Abuse & Neglect	**Grade retention**							
6762	Eckenrode et al., 1993		0.446	379	394	96.710	NaN	1.000	**0.446**
	Child Abuse & Neglect	**High school graduation**							
6729	Thornberry et al., 2001		-0.176	134	604	45.925	18.152	1.000	**-0.176**
6738	McGloin & Widom, 2001		-0.479	676	520	185.503	25.836	1.000	**-0.479**
6822	Lansford et al., 2007		-0.854	69	505	34.481	16.047	1.000	**-0.854**
6871	Boden et al., 2007		-0.158	171	800	64.465	20.480	1.000	**-0.158**
12770	Mersky & Topitzes, 2010		-0.407	179	1,148	99.452	23.057	1.000	**-0.407**

Record Id	Citation	Composite record Id	Unadjusted effect size	No. in test condition group	No. 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	1,923	42.612	25.745	1.000	**0.695**
12407	Thornberry et al., 2010		0.275	170	645	133.707	43.756	1.000	**0.275**
22339	Arteaga et al., 2010		0.273	117	1,091	43.577	26.094	1.000	**0.273**
6740	McGloin & Widom, 2001		0.135	676	520	193.432	48.674	1.000	**0.135**
22219	Huang et al., 2011		0.240	1,279	603	118.985	42.053	1.000	**0.240**
12949	Fergusson et al., 2008		0.113	162	839	38.871	24.330	1.000	**0.113**
	Child Abuse & Neglect	**Obesity**							
12420	Noll et al., 2007		0.543	84	89	23.508	21.636	1.000	**0.543**
22136	Power et al., 2015		0.100	766	3,373	322.370	147.427	1.000	**0.100**
22145	Power et al., 2015		-0.011	869	3,270	328.676	148.732	1.000	**-0.011**
22199	Bentley & Widom, 2009		0.004	410	303	174.235	106.153	1.000	**0.004**
22220	Shin & Miller, 2012		0.010	4,406	4,066	2114.582	240.737	1.000	**0.010**
	Child Abuse & Neglect	**PTSD**							
22262	Shenk et al., 2014		0.836	51	59	25.169	25.169	1.000	**0.836**
	Child Abuse & Neglect	**Special education**							
7488	Jonson-Reid et al., 2004		0.389	3,987	3,953	767.877	NaN	1.000	**0.389**
	Child Abuse & Neglect	**Test scores-academic**							
10750	Topitzes et al., 2010		-0.220	135	990	118.498	79.680	1.000	**-0.220**
6760	Eckenrode et al., 1993	22093	-0.383	206	206	101.147	71.440	1.000	**-0.383**
	Child Abuse & Neglect	**Tobacco (regular use)**							
12774	Mersky & Topitzes, 2010		0.387	143	919	65.624	NaN	1.000	**0.387**
	Crime (non-offender pop)	**High school graduation**							
28344	Hjalmarsson, 2008	28771	-0.304	1,222	6,195	634.275	36.690	1.000	**-0.304**
12777	Tanner et al., 1999		-0.403	478	1,882	130.899	30.013	1.000	**-0.403**
12823	Hirschfield, 2009		-0.666	216	2,039	31.603	17.445	1.000	**-0.666**
12930	Apel & Sweeten, 2009	28771	-0.623	400	4,649	233.531	33.376	1.000	**-0.623**
13721	Webbink et al., 2012		-0.595	224	2,028	99.944	28.023	1.000	**-0.595**
13722	Kirk & Sampson, 2009		-0.576	79	115	27.518	16.124	1.000	**-0.576**
	Crime (offender pop)	**High school graduation**							
28345	Hjalmarsson, 2008		-0.250	169	296	68.764	42.232	1.000	**-0.250**
28350	Apel & Sweeten, 2009		-0.079	656	1,036	199.486	70.677	1.000	**-0.079**
28351	Apel & Sweeten, 2009		-0.360	315	508	113.661	55.760	1.000	**-0.360**
30202	Hjalmarsson, 2008		-0.127	465	466	154.480	64.064	1.000	**-0.127**
	Depression	**High school graduation**							
	See Internalizing Composite								
	Depression	**Grade retention**							
	See Internalizing Composite								
	Diabetes	**Nursing home**							
21596	Valiyeva et al., 2006		0.714	173	3,353	47.770	33.920	1.000	**0.714**
21601	Valiyeva et al., 2006		0.246	255	2,681	89.900	50.837	1.000	**0.246**
21603	Harris & Cooper, 2006		0.213	21,148	116,484	15482.000	116.122	1.000	**0.213**
21607	Braunseis et al., 2011		0.079	460	1,841	54.080	36.985	1.000	**0.079**
21608	Luppa et al., 2010		-0.234	151	603	14.950	13.256	1.000	**-0.234**
21609	Stineman et al., 2012		0.285	928	6,908	157.810	67.187	1.000	**0.285**
21610	Andel et al., 2007		0.135	486	1,457	239.504	78.602	1.000	**0.135**
21634	Banaszak-Holl et al., 2004		0.110	901	5,775	210.220	75.165	1.000	**0.110**
	Disruptive Behavior	**Crime**							
	See Externalizing Composite								
	Disruptive Behavior	**Grade retention**							
12758	Galera et al., 2009		0.292	163	1,101	95.117	95.117	1.000	**0.292**
13148	Currie & Stabile, 2007		0.386	183	3,403	57.166	57.166	1.000	**0.386**
13169	Currie & Stabile, 2007		0.235	297	5,519	77.440	77.440	1.000	**0.235**
13755	Webbink et al., 2011		0.221	249	1,971	101.732	101.732	1.000	**0.221**
	Disruptive Behavior	**High school graduation**							
9847	Breslau et al., 2008	27134	-0.555	380	5,206	191.324	114.645	1.000	**-0.555**
8027	Fergusson & Lynskey, 1998		-0.333	83	886	41.020	35.875	1.000	**-0.333**
12757	Galera et al., 2009	36378	-0.438	71	643	20.951	19.522	1.000	**-0.438**
13710	Breslau et al., 2011	27127	-0.386	1,513	28,149	767.777	208.406	1.000	**-0.386**
13760	Porche et al., 2011		-0.525	287	2,245	129.576	89.180	1.000	**-0.525**
9853	Breslau et al., 2008		-0.389	704	4,882	288.496	200.000	1.000	**-0.389**
	Disruptive Behavior	**Special education**							
	See Externalizing Composite								
	Disruptive Behavior	**Test scores-academic**							
	See Externalizing Composite								
	Externalizing composite (includes conduct disorder & ADHD)	**Crime**							
12732	Fletcher & Wolfe, 2009		0.419	85	858	44.083	26.529	1.000	**0.419**
12742	Fergusson et al., 2005		0.763	46	927	17.420	13.809	1.000	**0.763**
12822	Murray et al., 2010		0.360	1,090	7,296	427.361	57.637	1.000	**0.360**
13152	Currie & Stabile, 2007	28580	0.192	164	3,056	105.543	40.841	1.000	**0.192**
13153	Currie & Stabile, 2007	28581	0.364	116	2,162	74.767	35.230	1.000	**0.364**
13159	Currie & Stabile, 2007	28580	0.101	323	2,903	197.645	49.826	1.000	**0.101**
13167	Currie & Stabile, 2007	28581	0.163	228	2,050	135.526	44.665	1.000	**0.163**
12735	Copeland et al., 2007		0.339	125	1,296	44.889	44.889	1.000	**0.339**
9481	Satterfield et al., 2007		0.535	169	64	25.454	25.454	1.000	**0.535**
13754	Webbink et al., 2011		0.501	98	778	26.655	26.655	1.000	**0.501**

Record Id	Citation	Composite record Id	Unadjusted effect size	No. in test condition group	No. in control group	Inverse variance weight - fixed effects	Inverse variance weight· random effects	WSIPP multipliers	Adjusted effect size
	Externalizing composite (includes conduct disorder & ADHD)	**High school graduation**							
612	McLeod & Kaiser, 2004		-0.273	57	367	22.039	22.039	1.000	**-0.273**
12008	Currie et al., 2010		-0.234	1,739	48,665	1090.491	1090.491	1.000	**-0.234**
13753	Webbink et al., 2011		-0.124	248	1,970	110.034	110.034	1.000	**-0.124**
	Externalizing composite (includes conduct disorder & ADHD)	**Special education**							
606	Fletcher & Wolfe, 2008		0.380	231	2,339	65.314	65.314	1.000	**0.380**
13149	Currie & Stabile, 2007	28630	0.254	139	2,592	39.018	39.018	1.000	**0.254**
13160	Currie & Stabile, 2007	28630	0.492	274	2,462	90.115	90.115	1.000	**0.492**
	Externalizing composite (includes conduct disorder & ADHD)	**Test scores-academic**							
12009	Currie et al., 2010		-0.168	1,739	48,665	1678.157	37.199	1.000	**-0.168**
27220	Currie & Stabile, 2007	28604	-0.272	258	2,318	230.977	32.663	1.000	**-0.272**
27221	Currie & Stabile, 2007	28604	-0.193	131	2,442	124.417	29.134	1.000	**-0.193**
13172	Currie & Stabile, 2007	28603	-0.202	121	2,256	114.931	28.582	1.000	**-0.202**
13173	Currie & Stabile, 2007	28603	-0.584	238	2,142	210.964	32.231	1.000	**-0.584**
12767	Massetti et al., 2008	27215	-0.019	85	130	51.393	32.359	1.000	**-0.019**
12768	Massetti et al., 2008	27215	-0.268	85	130	50.957	32.186	1.000	**-0.268**
27215	Turney & McLanahan, 2015	27235	0.044	821	1,481	527.110	39.900	1.000	**0.044**
	High school graduation	**Crime**							
7067	Lochner & Moretti, 2004		-0.183	102	102	50.788	50.244	1.000	**-0.183**
7069	Lochner & Moretti, 2004		-0.146	2,162	540	431.588	395.210	1.000	**-0.146**
12775	Ou & Reynolds, 2010		-0.211	374	359	119.628	116.651	1.000	**-0.211**
12795	Machin et al., 2011		-0.212	85	85	42.014	41.640	1.000	**-0.212**
13720	Webbink et al., 2012		-0.147	1,568	684	108.967	106.492	1.000	**-0.147**
13724	Bjerk, 2011		-0.293	1,286	672	437.130	399.851	1.000	**-0.293**
13740	Van Dorn et al., 2012		-0.158	28,987	5,666	527.975	474.539	1.000	**-0.158**
	Illicit drugs	**Crime**							
17708	Popovici et al., 2012		0.308	756	8,820	297.362	297.362	1.000	**0.308**
13743	Van Dorn et al., 2012		0.250	693	33,960	22.584	22.584	1.000	**0.250**
	Internalizing composite (includes depression & anxiety)	**Grade retention**							
13158	Currie & Stabile, 2007		0.313	640	2,958	150.643	150.643	1.000	**0.313**
13171	Currie & Stabile, 2007		0.234	1,038	4,793	224.204	224.204	1.000	**0.234**
	Internalizing composite (includes depression & anxiety)	**High school graduation**							
611	McLeod & Kaiser, 2004		-0.210	75	349	25.991	22.567	1.000	**-0.210**
1841	Duchesne et al., 2008		-0.215	177	1,640	93.564	60.514	1.000	**-0.215**
9850	Breslau et al., 2008	27133	-0.303	654	4,932	254.816	102.441	1.000	**-0.303**
9851	Breslau et al., 2008	27133	-0.159	1,782	3,804	455.939	124.524	1.000	**-0.159**
12011	Fletcher, 2010	28578	-0.167	186	2,141	55.172	41.732	1.000	**-0.167**
12937	Fergusson & Woodward, 2002		-0.058	124	840	43.426	34.644	1.000	**-0.058**
12951	Needham, 2009	28578	-0.140	1,566	12,666	365.892	116.681	1.000	**-0.140**
13716	Breslau et al., 2011	27128	0.012	978	28,684	357.238	115.787	1.000	**0.012**
13717	Breslau et al., 2011	27128	-0.059	6,025	23,637	1931.030	157.352	1.000	**-0.059**
13762	Porche et al., 2011		0.018	368	2,164	111.401	67.504	1.000	**0.018**
	Low birth weight (<2500 g)	**Infant mortality**							
34036	WSIPP (2017). WSIPP analysis to monetize birth indicators.		1.437	13,871	217,265	165.662	NaN	1.000	**1.437**
	Obesity	**Nursing home**							
21856	Elkins et al., 2006		0.159	917	4,367	776.944	776.944	1.000	**0.159**
21859	Valiyeva et al., 2006		0.196	645	2,881	103.024	103.024	1.000	**0.196**
31119	Valiyeva et al., 2006		0.224	537	2,399	246.284	246.284	1.000	**0.224**
	Opiods	**Crime**							
	See Illicit drugs								
	Preterm birth (<37 weeks gestation)	**Infant mortality**							
34037	WSIPP (2017). WSIPP analysis to monetize birth indicators.		1.103	19,270	204,015	193.365	193.365	1.000	**1.103**
	Small for gestational age	**Infant mortality**							
34039	WSIPP (2017). WSIPP analysis to monetize birth indicators.		0.794	14,462	208,580	165.558	165.558	1.000	**0.794**
	Smoking regularly < 14 years of age	**High school graduation**							
7080	Ellickson et al., 1998		-0.191	2,182	2,208	297.047	88.141	1.000	**-0.191**
7087	Bray et al., 2000		-0.345	926	466	30.053	24.240	1.000	**-0.345**
18257	Hawkins et al., 2013		-0.455	605	4,795	241.174	82.471	1.000	**-0.455**
18385	Yan & Brocksen, 2013		-0.321	490	1,323	74.566	46.751	1.000	**-0.321**
13718	Breslau et al., 2011		-0.411	9,818	19,844	3177.366	120.572	1.000	**-0.411**
	Smoking regularly < 18 years of age	**High school graduation**							
	See Smoking regularly < 14 years of age								
	Very low birthweight (<1500g)	**Infant mortality**							
34038	WSIPP (2017). WSIPP analysis to monetize birth indicators.		2.020	1,686	99,617	57.752	57.752	1.000	**2.020**

Exhibit A.I.3

Citations used in Linked Outcomes from Exhibits A.I.1 and A.I.2

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.

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.

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).

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.

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.

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., 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. (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).

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.

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.

Currie, J., & Stabile, M. (2007). *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.

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 (silver Spring, Md.), 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.

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. (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.

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.

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.

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*.

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.

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.

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, M., Luck, T., Matschinger, H.,, König, H-H., & Riedel-Heller, S.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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

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.

Turney, K., & McLanahan, S. (2015). The academic consequences of early childhood problem behaviors. *Social Science Research, 54*, 131-145.

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.

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 (2017). *WSIPP analysis to monetize birth indicators.*

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.

AII. Estimates of Human Capital Effects

Throughout this technical document we describe the parameters (derived from the research literature) we use to model the changes in labor market earnings from effects on human capital. Using the same meta-analytic approach we describe in Chapter 2, we take advantage of the research demonstrating the relationships between various outcomes and their impact on employment and/or earnings. Where sufficient research is available, we use meta-analysis to empirically estimate both the change in earnings and the change in the probability of employment given earnings due to a disorder or disease.

As described in Section 4.5d, we combine the results of these meta-analyses to compute an expected ratio: the ratio of total earnings for all people compared to the total earnings of a population with a specific disorder, condition, or experience. The mean changes in overall expected earnings are calculated based on the change in the probability of earning any money (i.e., the probability of employment) as well as the change in amount of earnings for those with earnings as a result of the disorder, condition, or experience. The ratio of total earnings (for both workers and non-workers) for individuals without the disorder to individuals with the condition is used in our modeling. The ratios are displayed in the relevant places in the technical document as follows:

- Substance abuse/dependence—See Exhibit 4.5.9
- Mental health disorders—See Exhibit 4.6.3
- Child abuse and neglect—See Exhibit 4.10.8
- Health conditions (obesity and diabetes)—See Exhibit 4.7.20

We list our current findings on these effects in the three Exhibits in this Appendix: Exhibit A.II.1 displays the meta-analytic results of each relationship we have estimated; Exhibit A.II.2 shows the individual studies for each effect; and Exhibit A.II.3 is a list of citations for all of the studies in these meta-analyses of human capital effects.

Exhibit A.II.1

Linked Outcomes Meta-Analytic Estimates of Standardized Mean Difference Effect Sizes

Estimated causal links between outcomes	No. 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
Alcohol disorder, leading to...											
Employment	2	-0.374	0.024	0.000	4.042	0.044	-0.360	0.051	0.000	**-0.360**	**0.051**
Earnings given employment	1	-0.051	0.040	0.204	0.000	0.000	-0.051	0.040	0.204	**-0.051**	**0.040**
Alcohol (problem use), leading to...											
Employment	6	-0.278	0.021	0.000	122.001	0.000	-0.204	0.105	0.052	**-0.204**	**0.105**
Earnings given employment	1	-0.019	0.045	0.677	0.000	0.000	-0.019	0.045	0.677	**-0.019**	**0.045**
Anxiety disorder, leading to...											
Employment	6	-0.165	0.290	0.000	27.176	0.000	-0.190	0.077	0.013	**-0.190**	**0.077**
Earnings given employment	2	-0.102	0.035	0.004	0.729	0.393	-0.102	0.035	0.004	**-0.102**	**0.035**
Child abuse & neglect, leading to...											
Employment	3	-0.247	0.075	0.001	2.754	0.252	-0.258	0.094	0.006	**-0.258**	**0.094**
Depression, leading to...											
Employment	11	-0.295	0.016	0.000	97.772	0.000	-0.336	0.065	0.000	**-0.336**	**0.065**
Earnings given employment	3	-0.022	0.021	0.284	1.083	0.582	-0.022	0.021	0.278	**-0.022**	**0.021**
Diabetes, leading to...											
Employment	5	-0.210	0.012	0.000	26.189	0.000	-0.252	0.043	0.000	**-0.252**	**0.043**
Earnings given employment	3	-0.027	0.030	0.366	0.417	0.812	-0.027	0.030	0.366	**-0.027**	**0.030**
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**
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**
PTSD, leading to...											
Employment	4	-0.391	0.022	0.000	26.311	0.000	-0.357	0.102	0.001	**-0.357**	**0.102**
Smoking regularly, leading to...											
Employment	5	-0.036	0.007	0.000	8.324	0.080	-0.045	0.012	0.000	**-0.045**	**0.012**
Earnings given employment	4	-0.056	0.006	0.000	11.809	0.008	-0.054	0.020	0.008	**-0.054**	**0.020**

Exhibit A.II.2

Linked Outcomes

Individual Estimates of Standardized Mean Difference Effect Sizes

Record Id	Citation	Composite record Id	Un-adjusted ES	No. in test condition group	No. in control group	Inverse variance weight - fixed effects	Inverse variance weight-random effects	WSIPP multipliers	Adjusted ES
	Alcohol disorder				**Employment**				
7118	Mullahy & Sindelar, 1996		-0.406	2,232	21,573	1,224.110	115.083	1.000	**-0.406**
13756	French et al., 2011	30766	-0.211	1,910	18,459	481.485	100.509	1.000	**-0.211**
13773	Sangchai, 2006	30766	-0.371	1,689	16,339	658.965	106.496	1.000	**-0.371**
	Alcohol disorder				**Earnings given employment**				
9644	Jones & Richmond, 2006		-0.051	798	2,848	622.899	NaN	1.000	**-0.051**
	Alcohol (problem use)				**Employment**				
35619	Jorgensen et al., 2017		-0.033	5,334	12,356	452.574	8.906	1.000	**-0.033**
10597	Saffer & Dave, 2005		-0.082	210	6,790	125.932	8.474	1.000	**-0.082**
7117	Mullahy & Sindelar, 1996	7119	-0.412	2,381	21,425	1,283.028	9.021	1.000	**-0.412**
7122	Terza, 2002	7119	-1.042	982	8,840	487.473	8.919	1.000	**-1.042**
7161	Feng et al., 2001		0.028	647	7,475	245.544	8.761	1.000	**0.028**
7163	Auld (2002)	7119	-0.602	982	8,840	387.087	8.877	1.000	**-0.602**
7167	MacDonald & Shields, 2004		-0.217	664	5,980	298.821	8.817	1.000	**-0.217**
13757	French et al., 2011		-0.312	1,910	18,459	534.719	8.933	1.000	**-0.312**
	Alcohol (problem use)				**Earnings given employment**				
12355	Keng & Huffman, 2010	30885	-0.019	1,393	2,707	919.795	919.795	1.000	**-0.019**
7178	Bray, 2005	30885	-0.017	277	1,572	235.743	235.743	1.000	**-0.017**
	Anxiety disorder				**Employment**				
9633	Cornwell et al., 2009		-0.101	1,128	9,513	139.938	44.364	1.000	**-0.101**
11798	Gibb et al., 2010		-0.151	143	808	38.964	24.355	1.000	**-0.151**
12127	Cowell et al., 2009	30947	-0.100	2,301	30,774	1,388.209	62.054	1.000	**-0.100**
12345	Chatterji et al., 2009		-0.117	1,168	10,645	561.795	58.225	1.000	**-0.117**
12349	Baldwin et al., 2007		-0.583	294	9,675	136.104	43.971	1.000	**-0.583**
7185	Ettner et al., 1997		-0.087	562	4,064	163.088	46.455	1.000	**-0.087**
17732	Burnett-Zeigler et al., 2013	30947	-0.174	139	22,268	52.655	29.081	1.000	**-0.174**
	Anxiety disorder				**Earnings given employment**				
12352	Baldwin et al., 2007		-0.143	294	10,530	285.937	201.543	1.000	**-0.143**
7188	Ettner et al., 1997	30906	-0.029	562	4,064	493.823	286.577	1.000	**-0.029**
7197	Marcotte & Wilcox-Gők, 2003	30906	-0.123	752	2,679	586.246	315.435	1.000	**-0.123**
	Child abuse & neglect				**Employment**				
23093	Mersky & Topitzes, 2010		-0.184	184	1,178	107.541	58.794	1.000	**-0.184**
21862	Currie & Widom, 2010		-0.470	174	174	42.222	31.853	1.000	**-0.470**
22263	Covey et al., 2013		-0.156	124	1,169	27.583	22.746	1.000	**-0.156**
	Depression				**Employment**				
9632	Cornwell et al., 2009		-0.401	724	9,917	60.495	22.539	1.000	**-0.401**
11797	Gibb et al., 2010		-0.351	143	808	46.621	20.289	1.000	**-0.351**
12126	Cowell et al., 2009	30948	-0.298	1,534	31,541	989.825	34.665	1.000	**-0.298**
12344	Chatterji et al., 2009		-0.310	1,709	10,104	861.651	34.485	1.000	**-0.310**
12346	Tian et al., 2005		-0.150	459	5,239	279.998	31.838	1.000	**-0.150**
12348	Baldwin et al., 2007		-0.690	703	10,121	328.193	32.379	1.000	**-0.690**
7184	Ettner et al., 1997		-0.328	454	4,172	170.633	29.675	1.000	**-0.328**
7193	Farahati et al., 2003		-0.255	74	438	32.841	17.156	1.000	**-0.255**
7202	Savoca & Rosenheck, 2000		-0.315	79	1,338	31.662	16.829	1.000	**-0.315**
7216	Alexandre & French, 2001		-0.527	384	890	144.464	28.769	1.000	**-0.527**
17730	Burnett-Zeigler et al., 2013	30948	-0.234	350	22,057	139.380	28.561	1.000	**-0.234**
17799	Peng et al., 2013		-0.081	1,386	13,841	681.701	34.124	1.000	**-0.081**

Record Id	Citation	Composite record Id	Un-adjusted ES	No. in test condition group	No. in control group	Inverse variance weight - fixed effects	Inverse variance weight-random effects	WSIPP multipliers	Adjusted ES
	Depression				**Earnings given employment**				
12351	Baldwin et al., 2007		-0.054	703	10,121	657.283	487.697	1.000	**-0.054**
7187	Ettner et al., 1997	30905	-0.086	454	4,172	409.028	336.263	1.000	**-0.086**
7196	Marcotte & Wilcox-Gők, 2003	30905	0.032	483	2,948	415.199	340.423	1.000	**0.032**
17800	Peng et al., 2013		-0.004	1,386	13,841	1,259.512	755.859	1.000	**-0.004**
	Diabetes				**Employment**				
21545	Ng et al., 2001	31015	-0.035	1,351	67,283	658.291	60.884	1.000	**-0.035**
21546	Tunceli et al., 2005	31016	-0.258	490	6,565	186.008	49.306	1.000	**-0.258**
21552	Stewart et al., 2007		-0.458	1,033	18,042	167.206	47.878	1.000	**-0.458**
21577	Minor, 2013		-0.194	7,808	95,780	4,453.542	66.093	1.000	**-0.194**
31013	Kahn, 1998	31016	-0.373	959	8,738	577.691	60.108	1.000	**-0.373**
31014	Kahn, 1998	31015	-0.352	2,352	87,548	1,522.654	64.258	1.000	**-0.352**
33922	WSIPP, 2017		-0.092	1,085	12,382	514.140	59.345	1.000	**-0.092**
	Diabetes				**Earnings given employment**				
21578	Minor, 2013		-0.065	221	32,302	219.413	219.413	1.000	**-0.065**
21584	Songer et al., 1989		0.000	127	127	63.500	63.500	1.000	**0.000**
21591	Kahn, 1998		-0.019	959	8,738	864.144	864.144	1.000	**-0.019**
	Drug disorder				**Employment**				
7105	Zuvekas et al., 2005		-0.171	929	8,089	226.909	62.062	1.000	**-0.171**
7169	Alexandre & French, 2004		-0.285	926	553	226.149	62.005	1.000	**-0.285**
7171	French et al., 2001		-0.271	379	9,242	215.995	61.216	1.000	**-0.271**
7190	Ettner et al., 1997		-0.624	148	4,478	78.022	40.779	1.000	**-0.624**
5574	Buchmueller & Zuvekas, 1998		-0.220	449	1,651	178.804	57.808	1.000	**-0.220**
	Obesity				**Employment**				
21813	Han et al., 2009		-0.023	16,305	95,924	5,624.168	146.322	1.000	**-0.023**
21825	Tunceli et al., 2006		-0.156	526	2,419	232.355	91.239	1.000	**-0.156**
	Obesity				**Earnings given employment**				
30992	Dastan, 2011	31012	-0.022	4,037	8,069	2,690.748	2,690.748	1.000	**-0.022**
31004	Baum et al., 2006	31012	-0.045	1,462	3,761	1,052.746	1,052.746	1.000	**-0.045**
	PTSD				**Employment**				
7207	Savoca & Rosenheck, 2000	30945	-0.374	315	1,102	100.821	23.941	1.000	**-0.374**
17985	WSIPP, 2013		-0.440	2,496	32,157	1,554.162	30.775	1.000	**-0.440**
18013	Resnick & SG, 2008		-0.128	925	4,901	317.868	28.574	1.000	**-0.128**
18051	McCarren et al., 1995		-0.432	273	273	46.575	18.754	1.000	**-0.432**
18054	Zatzick et al., 1997	30945	-0.723	242	948	43.681	18.267	1.000	**-0.723**
	Smoking regularly				**Employment**				
36437	Strong et al., 2014		-0.120	199	199	58.855	57.737	1.000	**-0.120**
12807	Jofre-Bonet et al., 2005		-0.020	31,105	88,778	12,120.635	2,428.834	1.000	**-0.020**
12819	Dastan, 2011		-0.073	4,005	8,004	1,330.124	925.047	1.000	**-0.073**
13188	WSIPP, 2014		-0.047	11,082	26,357	4,849.319	1,867.655	1.000	**-0.047**
13189	WSIPP, 2014		-0.062	9,064	22,850	3,234.449	1,566.445	1.000	**-0.062**
	Smoking regularly				**Earnings given employment**				
9658	Anger & Kvasnicka, 2010		-0.164	819	1,149	476.614	359.808	1.000	**-0.164**
18497	Baum et al., 2006	30943	-0.019	1,462	3,761	1,052.919	613.171	1.000	**-0.019**
18498	Cowan & Schwab, (2011).	30943	-0.042	1,903	4,237	1,313.093	693.151	1.000	**-0.042**
12808	Jofre-Bonet et al., 2005		-0.061	31,105	88,778	23,026.419	1,380.157	1.000	**-0.061**
12820	Dastan, 2011	30943	-0.029	4,037	8,069	2,690.646	949.862	1.000	**-0.029**
12931	Braakmann, 2008		-0.013	3,611	8,647	2,547.185	931.344	1.000	**-0.013**

Exhibit A.II.3

Citations used in Linked Outcomes from Exhibits A.II.1 and AI.I.2

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.

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.

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.

Baldwin, M.L., & Marcus, S.C. (2007). Labor market outcomes of persons with mental disorders. *Industrial Relations, 46*(3), 481-510.

Baum, C.L., Ford, W.F., & Hopper, J.D. (2006). The obese smoker's wage penalty. *Social Science Quarterly, 87*(4), 863-881.

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.

Bray, J.W. (2005). Alcohol use, human capital, and wages. *Journal of Labor Economics, 23*(2), 279-312.

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.

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.

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., & Widom, C.S. (2010). Long-term consequences of child abuse and neglect on adult economic well-being. *Child Maltreatment, 15*(2), 111-120.

Dastan, I. (2011). Labor market effects of obesity, smoking, and alcohol use. *Dissertation Abstracts International, 72*(03), A.

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.

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.

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.

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.

Han, E., Norton, E.C., & Stearns, S.C. (2009). Weight and wages: fat versus lean paychecks. *Health Economics, 18*(5), 535-548.

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.

Jorgensen, M.B., Thygesen, L. C., Becker, U., & Tolstrup, J.S. (2017). Alcohol consumption and risk of unemployment, sickness absence and disability pension in Denmark: a prospective cohort study: Alcohol consumption and unemployment. *Addiction*.

Kahn, M.E. (1998). Health and Labor Market Performance: The Case of Diabetes. *Journal of Labor Economics, 16*(4), 878-899.

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.

MacDonald, Z., & Shields, M.A. (2004). Does problem drinking affect employment? Evidence from England. *Health Economics, 13*(2), 139-155.

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.

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.

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.

Mullahy, J., & Sindelar, J.L. (1996). Employment, unemployment, and problem drinking. *Journal of Health Economics, 15*(4), 409-434.

Ng, Y.C., Jacobs, P., & Johnson, J.A. (2001). Productivity losses associated with diabetes in the US. *Diabetes Care, 24*(2), 257-61.

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.

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.

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.

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.

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.

Strong, C., Juon, H.-S., & Ensminger, M.E. (2014). Long-term effects of adolescent smoking on depression and socioeconomic status in adulthood in an Urban African American cohort. *Journal of Urban Health, 91*(3), 526-540.

Terza, J.V. (2002). Alcohol abuse and employment: A second look. *Journal of Applied Econometrics, 17*(4), 393-404.

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.

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.

WSIPP (2017). *WSIPP analysis of diabetes' effect on employment.* Olympia: Washington State Institute for Public Policy.

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.

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.

AIII. Regression Results

The below Exhibits contain the full regression results referenced throughout the Technical Document.

Exhibit A.III.1

Two-Part Model Assessing Health Care 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)
	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	***	(0.90 - 1.16)
Flu shot		0.80	***	(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	0.50		(-0.28 - 1.28)
	Agree somewhat/strongly	0.14		(-0.66 - 0.93)
Smoke history		**0.06**		**(-0.08 - 0.20)**
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	0.10		(-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	0.10		(-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	0.00		(-0.10 - 0.09)
	Divorced, separated, widowed	0.09	**	(0.00 - 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 - -0.10)
	High income (GE 400%)	-0.20	***	(-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 - 0.50)
	Agree somewhat/strongly	-0.38	*	(-0.79 - 0.02)
Smoke history		**0.25**	*******	**(0.17 - 0.32)**
Intercept		8.28	***	(7.79 - 8.78)

Notes:
Significance levels: * p<0.1, ** p<0.05, *** p<0.001.
Number of observations = 17,899
Weighted size = 513,466,894
Design df = 204
F(30, 175) = 55.98
Prob > F = 0.0000

Exhibit A.III.2

Two-Part Model Assessing Health Care 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)

Notes:
Significance levels: * p<0.1, ** p<0.05, *** p<0.001.
Number of observations = 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

Exhibit A.III.3

Two-Part Model Assessing Non-Treatment Health Care 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)

Notes:
Significance levels: * p<0.1, ** p<0.05, *** p<0.001.
Number of observations = 5,522
Weighted size = 229,038,154
Design df = 200
F(20, 181) = 31.78
Prob > F =0.0000

Exhibit A.III.4

Two-Part Model Assessing Non-Treatment Health Care 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)

Notes:
Significance levels: * p<0.1, ** p<0.05, *** p<0.001.
Number of observations = 5,522
Weighted. size = 229,038,154
Design df = 200
F(14, 187) = 35.27
Prob > F = 0.0000

Exhibit A.III.5

Two-Part Model Assessing Non-Treatment Health Care Costs of Child Emotional Conditions (Depression, Anxiety)

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age		-0.04	***	(-0.06, -0.01)
Female		0.40	***	(0.16, 0.65)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.67	***	(-0.99, -0.35)
	Black, non-Hispanic	-0.89	***	(-1.21, -0.56)
	Asian non-Hispanic	-0.71	**	(-1.39, -0.02)
Poverty status (ref: High income)	Poor	-0.53	***	(-081, -0.26)
Region (ref: West)	Northeast	0.39	*	(-0.03, 0.82)
	Midwest	0.68	***	(0.28, 1.07)
	South	0.48	***	(0.16, 0.81)
Urban-Rural	MSA	0.50	**	(0.09, 0.91)
Chronic conditions	Asthma	1.06	***	(0.53, 1.60)
	Other^	0.19		(-0.74, 1.12)
Uninsured		-0.99	***	(-1.27, -0.70)
Emotional condition indicated (SDQ)		**0.24**		(-0.30, 0.78)
Intercept		1.68	***	(1.13, 2.23)
Part two: GLM, estimated costs				
Age		0.02		(-0.04, 0.07)
Female		0.13		(-0.25, 0.50)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.26		(-0.81, 0.30)
	Black, non-Hispanic	-0.61	***	(-0.87, -0.34)
	Asian non-Hispanic	-0.55	*	(-1.21, 0.11)
Poverty status (ref: High income)	Poor	-0.23		(-0.57, 0.10)
Region (ref: West)	Northeast	-0.09		(-0.81, 0.63)
	Midwest	-0.19		(-0.88, 0.50)
	South	-0.31		(-0.93, 0.30)
Urban-Rural	MSA	0.07		(-0.27, 0.42)
Chronic conditions	Asthma	0.39	**	(0.08, 0.70)
	Other*	0.96	***	(0.28, 1.64)
Uninsured		-0.16		(-0.70, 0.38)
Emotional condition indicated (SDQ)		**0.52**	**	(0.00, 1.04)
Intercept		7.20	***	(6.07, 8.33)

Notes:

^Conditions include Down syndrome, cerebral palsy, muscular dystrophy, cystic fibrosis, sickle cell anemia, arthritis, congenital heart disease, and other heart disease.

Significance levels: * p<0.1, ** p<0.05, *** p<0.001.

Number of observations = 3,133

Weighted size = 26,229,116

Design df = 249

F(14, 236) = 12.15

Prob > F = 0.0000

Exhibit A.III.6

Two-Part Model Assessing Non-Treatment Health Care Costs of Child Conduct Condition (Disruptive Behavior)

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age		-0.03	**	(-0.06, -0.01)
Female		0.40	***	(0.16, 0.64)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.69	***	(-1.00, -0.38)
	Black, non-Hispanic	-0.92	***	(-1.24, -0.60)
	Asian non-Hispanic	-0.72	**	(-1.39, -0.04)
Poverty status (ref: High income)	Poor	-0.52	***	(-0.78, -0.25)
Region (ref: West)	Northeast	0.37	*	(-0.06, 0.80)
	Midwest	0.66	***	(0.27, 1.06)
	South	0.50	***	(0.18, 0.82)
Urban-rural	MSA	0.52	***	(0.12, 0.93)
Chronic conditions	Asthma	1.07	***	(0.53, 1.60)
	Other^	0.16		(-0.77, 1.09)
Uninsured		-1.01	***	(-1.30, -0.71)
Conduct condition indicated (SDQ)		**0.40**	*	(-0.05, 0.85)
Intercept		1.63	***	(1.06, 2.19)
Part two: GLM, estimated costs				
Age		0.02		(-0.03, 0.08)
Female		0.22		(-0.13, 0.58)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.30		(-0.85, 0.26)
	Black, non-Hispanic	-0.69	***	(-0.96, -0.42)
	Asian non-Hispanic	-0.57	*	(-1.22, 0.07)
Poverty status (ref: High income)	Poor	-0.31	*	(-0.65, 0.03)
Region (ref: West)	Northeast	-0.15		(-0.83, 0.53)
	Midwest	-0.23		(-0.89, 0.43)
	South	-0.41		(-1.02, 0.19)
Urban-rural	MSA	0.08		(-0.27, 0.43)
Chronic conditions	Asthma	0.46	***	(0.13, 0.79)
	Other*	0.88	***	(0.29, 1.47)
Uninsured		-0.16		(-0.70,0.37)
Conduct condition indicated (SDQ)		**0.88**	***	(0.46, 1.31)
Intercept		7.03	***	(5.87, 8.19)

Notes:

^Conditions include Down syndrome, cerebral palsy, muscular dystrophy, cystic fibrosis, sickle cell anemia, arthritis, congenital heart disease, and other heart disease.

Significance levels: * p<0.1, ** p<0.05, *** p<0.001.

Number of observations = 3,132

Weighted size = 26,203,162

Design df = 249

F(14, 236) = 11.54

Exhibit A.III.7

Two-Part Model Assessing Non-Treatment Health Care Costs of Child ADHD

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age		-0.02		(-0.04, 0.00)
Female		0.02		(-0.18, 0.21)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.34	**	(-0.64, -0.05)
	Black, non-Hispanic	-0.65	***	(-0.97, -0.34)
	Asian non-Hispanic	-0.68	***	(-1.06, -0.20)
	Other	-0.42		(-1.02, 0.17)
Income status (ref: High income)	Poor	-0.89	***	(-1.30, -0.49)
	Low income	-0.74	***	(-1.14, -0.35)
	Middle income	-0.53	***	(-0.91, -0.15)
Region (ref: West)	Northeast	0.09		(-0.28, 0.47)
	Midwest	0.02		(-0.34, 0.38)
	South	0.01		(-0.30, 0.33)
Chronic conditions	Asthma	0.82	***	(0.41, 1.24)
	Other^	1.17	**	(0.11, 2.24)
Uninsured		-1.17	***	(-1.51, -0.84)
ADHD Diagnosis		**0.64**	***	(0.24, 1.04)
Intercept		2.76	***	(2.23, 3.28)
Part two: GLM, estimated costs				
Age		0.04	***	(0.01, 0.06)
Female		0.04		(-0.11, 0.20)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.19	*	(-0.39, 0.01)
	Black, non-Hispanic	-0.49	***	(-0.75, -0.23)
	Asian non-Hispanic	-0.27	**	(-0.54, -0.01)
	Other	0.20		(-0.13, 0.53)
Income status (ref: High income)	Poor	-0.48	***	(-0.73, -0.22)
	Low income	-0.46	***	(-0.71, -0.22)
	Middle income	-0.31	***	(-0.55, -0.07)
Region (ref: West)	Northeast	0.30	***	(0.07, 0.53)
	Midwest	0.18		(-0.06, 0.41)
	South	-0.17	*	(-0.36, 0.02)
Chronic conditions	Asthma	0.51	***	(0.28, 0.74)
	Other*	1.77	***	(1.46, 20.8)
Uninsured		-0.78	***	(-1.19, -0.36)
ADHD Diagnosis		**0.52**	***	(0.25, 0.78)
Intercept		7.09	***	(6.72, 7.47)

Notes:

^Conditions include diabetes, paralysis, epilepsy, heart disease.

Significance levels: * p<0.1, ** p<0.05, *** p<0.001.

Number of observations = 6,945

Weighted size = 54,149,172

Design df = 204

F(16, 189) = 5.52

Prob > F = 0.0000

Exhibit A.III.8

Two-Part Model Assessing Non-Treatment Health Care Costs of Child Internalizing Condition

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age		-0.03	**	(-0.06, -0.01)
Female		0.40	***	(0.16, 0.64)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.69	***	(-1.00, -0.38)
	Black, non-Hispanic	-0.92	***	(-1.24, -0.60)
	Asian non-Hispanic	-0.74	**	(-1.42, -0.06)
Poverty status (ref: High income)	Poor	-0.47	***	(-0.74, -0.21)
Region (ref: West)	Northeast	0.40	*	(-0.02, 0.82)
	Midwest	0.67	***	(0.27, 1.06)
	South	0.51	***	(0.19, 0.82)
Urban-rural	MSA	0.51	**	(0.10, 0.91)
Chronic conditions	Asthma	1.07	***	(0.53, 1.60)
	Other^	0.23		(-0.69, 1.14)
Uninsured		-1.00	***	(-1.29, -0.71)
Internalizing condition indicated (SDQ)		**-0.14**		(-0.57, 0.29)
Intercept		1.67	***	(1.12, 2.23)
Part two: GLM, estimated costs				
Age		0.02		(-0.04, 0.07)
Female		0.14		(-0.22, 0.50)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.26		(-0.81, 0.29)
	Black, non-Hispanic	-0.60	***	(-0.87, -0.34)
	Asian non-Hispanic	-0.58	*	(-1.23, 0.07)
Poverty status (ref: High income)	Poor	-0.24		(-0.57, 0.09)
Region (ref: West)	Northeast	-0.13		(-0.84, 0.59)
	Midwest	-0.21		(-0.89, 0.47)
	South	-0.34		(-0.95, 0.28)
Urban-rural	MSA	0.07		(-0.27, 0.42)
Chronic conditions	Asthma	0.39	**	(0.05, 0.73)
	Other*	0.95	***	(0.26, 1.63)
Uninsured		-0.15		(-0.69, 0.39)
Internalizing condition indicated (SDQ)		**0.43**	**	(0.07, 0.79)
Intercept		7.22	***	(6.10, 8.34)

Notes:

^Conditions include Down syndrome, cerebral palsy, muscular dystrophy, cystic fibrosis, sickle cell anemia, arthritis, congenital heart disease, and other heart disease.

Significance levels: * p<0.1, ** p<0.05, *** p<0.001.

Number of observations = 3,132

Weighted size = 26,203,162

Design df = 249

F(14, 236) = 10.96

Prob > F = 0.0000

Exhibit A.III.9

Two-Part Model Assessing Non-Treatment Health Care Costs of Child Externalizing Condition

Category	Variable	Coefficient		95% CI
Part one: Logit, probability of incurring costs				
Age		-0.03	**	(-0.06, -0.01)
Female		0.41	***	(0.16, 0.65)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.68	***	(-0.99, -0.37)
	Black, non-Hispanic	-0.92	***	(-1.24, -0.59)
	Asian non-Hispanic	-0.74	**	(-1.42, -0.05)
Poverty status (ref: High income)	Poor	-0.49	***	(-0.75, -0.23)
Region (ref: West)	Northeast	0.38	*	(-0.05, 0.81)
	Midwest	0.66	***	(0.27, 1.05)
	South	0.50	***	(0.19, 0.82)
Urban-rural	MSA	0.51	***	(0.10, 0.92)
Chronic conditions	Asthma	1.06	***	(0.53, 1.59)
	Other^	0.20		(-0.73, 1.12)
Uninsured		-1.00	***	(-1.29, -0.71)
Externalizing condition indicated (SDQ)		**0.12**		(-0.26, 0.50)
Intercept		1.64	***	(1.06, 2.21)
Part two: GLM, estimated costs				
Age		0.03		(-0.03, 0.08)
Female		0.24		(-0.11, 0.59)
Race/ethnicity (ref: White, non-Hispanic)	Hispanic	-0.27		(-0.84, 0.30)
	Black, non-Hispanic	-0.68	***	(-0.95, -0.41)
	Asian non-Hispanic	-0.57	*	(-1.23, 0.08)
Poverty status (ref: High income)	Poor	-0.28	*	(-0.62, 0.06)
Region (ref: West)	Northeast	-0.13		(-0.84, 0.58)
	Midwest	-0.22		(-0.89, 0.45)
	South	-0.42		(-1.02, 0.18)
Urban-rural	MSA	0.12		(-0.24, 0.47)
Chronic conditions	Asthma	0.44	***	(0.11, 0.77)
	Other*	0.90	***	(0.21, 1.59)
Uninsured		-0.14		(-0.68, 0.39)
Externalizing condition indicated (SDQ)		**0.64**	***	(0.26, 1.01)
Intercept		6.97	***	(5.80, 8.15)

Notes:

^Conditions include Down syndrome, cerebral palsy, muscular dystrophy, cystic fibrosis, sickle cell anemia, arthritis, congenital heart disease, and other heart disease.

Significance levels: * p<0.1, ** p<0.05, *** p<0.001.

Number of observations = 3,131

Weighted size = 26,192,011

Design df = 249

F(14, 236) = 11.15

Prob > F = 0.0000

Exhibit A.III.10

Two-Part Model Assessing Health Care 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)

Notes:

Significance levels: * p<0.1, ** p<0.05, *** p<0.001.
Number of observations = 38,974
Weighted size = 313,489,853
Design df = 203
F(15,189) = 170.30
Prob > F = 0.0000

For further information, contact:
Michael Hirsch at 360.664.9081, michael.hirsch@wsipp.wa.gov