2024

CONSTELLATION FUND

TECHNICAL DOCUMENTATION

FINDING THE STARS AND CONNECTING THE DOTS IN THE FIGHT AGAINST POVERTY IN THE TWIN CITIES

WWW.CONSTELLATIONFUND.ORG

1. Introduction

The Constellation Fund, created in 2018, supports organizations in the seven-county Twin Cities area that provide services to raise the living standards of people experiencing poverty while also yielding demonstrable results. Constellation is evidence-directed, heart-driven, and community-informed.

This report describes the technical approach underlying the evaluation methods Constellation Fund uses to help guide program investments. They were created under the guidance of the Constellation Impact Council (CIC) by adapting and updating the metrics framework, formulas, and references developed by the Robin Hood Foundation to fit the seven-county Twin Cities metropolitan area. The CIC is chaired by Aaron Sojourner, Senior Economist at the W.E. Upjohn Institute for Employment Research, and includes Judy Temple, Professor at the Humphrey School of Public Affairs, University of Minnesota, and Abigail Wozniak, Director of the Opportunity & Inclusive Growth Institute at the Federal Reserve Bank of Minneapolis.

With the oversight of the CIC, Constellation will continue to develop its methods, which will be updated and reported on regularly in this document. For additional information about Constellation's evidence-driven approach to poverty alleviation, please visit ConstellationFund.org. For the list of current metrics, see Constellation's Metrics.

2. Overview of Constellation's Benefit-Cost Framework

2.1. General Equations

Constellation uses benefit-cost analysis (BCA) to assess the economic impact of poverty-fighting interventions. BCA consists of comparing the benefits of a program to the cost of delivering it. The result is a number we call the Benefit-Cost Ratio (BCR), which indicates the dollars in benefits generated by the program for every dollar invested in it. Constellation's BCA is implemented from the perspective of participants in the programs. The only benefits included in the estimations are those accrued by participants and the only costs used are program costs. Participants' benefits include increased income, improvements in health, and similar impacts. Our estimations contrast the notion of social return on investment (ROI) analysis where benefits and costs accrued by other agents in society are included, for example, taxpayers, non-participant individuals, or other institutions. Social ROI estimates are useful in many policy and analytical frameworks; however, Constellation seeks to find interventions that improve the lives of those living in poverty. Societal benefits generated directly or indirectly by Constellation's grantees include taxpayers and other sectors who are not disadvantaged or living in poverty.

$$BCR = \frac{Private\ benefits}{Program\ Costs}$$

Constellation uses BCRs to rank grant applicants and complements the economic assessment with qualitative evaluation to make investment recommendations to the Board of Directors.

Constellation estimates the expected benefits of a proposed program investment based on a general model of the value of the stream of expected future benefits for low-income individuals and families.

We estimate these benefits by using a system of equations we refer to as "Metrics." The general model of Constellation's metrics is depicted as:

$$Private\ benefits\ =\ N\times\sum_{t=1}^{T}\frac{Q_t\times P_t}{(1+Dis)^t}$$

Q_t is defined as units of change in the outcome of interest, usually measured as the average per-participant impact from the proposed program in the year (*t*) years after the start of the program. It is the difference in the average outcome among potential participants between two possibilities:

1) if all participants receive access to the program, or
2) if none do

Program impacts are estimated based on evidence from evaluation results of individual programs or average effect size from several evaluations of comparable interventions.

P_t is our best estimate of the monetary value of a unit change in the outcome Q_t at year *t*.

D is the social discount rate, which we establish as 3 percent. This adjustment reflects the fact that a dollar today is worth more than a dollar in the future.

T is the number of years that any program effects are expected to last. This is estimated based on a combination of factors, including evidence from research literature and the assessment of the program made by Constellation's impact officers.

N is the projected number of participants.

Computing Method Examples

Example of Q estimated using program's outcome data:

For a proposed job search assistance program, we expect that the program will have an effect on the participants' employment for two years after participation. Q in year one would be the average expected impact on participants' employment probability in the first year after finishing the program. Suppose the program assists 100 participants. If 90 of those 100 are employed at the end of the year and only 50 of those 100 would have found employment if they did not receive assistance from the program, then Q_1 would equal a 40 percentage point increase in participant's employment. In year two after the program, if 85 participants are employed, but only 60 would be employed in absence of the program, then Q_2 equals 25 percentage points. In subsequent years let's assume there is no difference in employment probabilities between participants and non-participants.

P would equal the average earnings among employed participants at (*t*) years post-program. Suppose this is $15,000 per year, and let's assume a 3 percent discount rate. Then, the value of the program's benefits would be:

$$Private\ benefits\ =\ 100\times\frac{50\%\times\$15,000}{(1+3\%)^1}\ +\ 100\times\frac{25\%\times\$15,000}{(1+3\%)^2}\ =\ \$10,816$$

If the per-participant proposed program cost $4,000 to operate each year, and participants require a full year of of services to ensure the increased chances of employment in the future, the Benefit Cost Ratio of the program is:

$$BCR = \frac{\$10{,}816}{\$4{,}.000} = \$2.70$$

The BCR indicates that participants of this program are expected to receive $2.70 per $1 invested in the program. Constellation uses this ratio to compare programs and find the interventions with the highest returns.

Example of Q estimated using evidence from research studies:

Suppose that we are evaluating the same job assistance organization mentioned in the example above. However, the organization does not collect outcome data from its participants, so the number of people actually employed after their participation is unknown. In this case, Constellation would research the literature on workforce programs and look for evidence from similar programs that have demonstrated effectiveness. Let's assume that from our studies, we learned that the average impact of these interventions on the employment rate is a 40% increase after one year. If the rate of employment of non-participants is 50%, the Q value would be 20 percentage points (a 40% increase from 50%), and participants are expected to have a 70% chance of finding a job. Suppose that the studies didn't track the impact of the program beyond year one. Then, Constellation would estimate benefits only for the timeframe observed in the research. In this case the benefits would be:

$$Private\ benefits = 100 \times \frac{20\% \times \$15{,}000}{(1+3\%)^1} = \$2,913$$

The BCR would be:

$$BCR = \frac{\$2{,}918}{\$4{,}000} = \$0.73$$

In this scenario, Constellation wouldn't fund the program since the costs exceed the private benefits generated by it. Note that these two examples are designed for illustrative purposes and do not imply that BCRs estimated using program data are higher than those estimated using researched evidence.

2.2. General Characteristics of Constellation's Benefit-Cost Model

Internally Consistent Estimates

Constellation develops each BCR using a common set of assumptions, methods, and parameters with the objective to conduct comparisons across programs with different delivery methods and outcomes. The BCR results and any comparisons across organizations are valid only for the set of programs evaluated by Constellation, and should not be generalized beyond this sample.

We acknowledge that the Benefit-Cost estimates are sensitive to assumptions and may include idiosyncratic errors and biases that could affect some programs more than others. Constellation makes an effort to minimize these biases using several quality checks throughout the assessment process. We document critical assumptions in the Strength of Evidence section in each metric. We also share with grantees all methods, assumptions, and data used for each metric.

General Assumptions

These assumptions are relevant in the application of all metrics.

- Constellation calculates private monetary benefits for program participants, generally for individuals at least 200% the federal poverty guideline.

- All monetary values are converted to constant dollars using the U.S. Price Indexes for Personal Consumption Expenditures from the U.S. Bureau of Economic Analysis, "Table 2.4.4U. Price Indexes for Personal Consumption Expenditures by Type of Product.

- Present values are calculated using a social discount rate of 3% of future values. This reflects a higher value of benefits today than benefits in the future, but is on the lower end of common social discount rates (2 – 10%), hence future benefits receive a relatively heavy weight.

- Whenever appropriate, earnings are reported net of taxes to assess the private benefit accrued by participants. We update these sources every three years.

- We use the best available evidence to estimate each program's impact on income and health outcomes of participants. This evidence may come from scientific research in peer-reviewed journals, private program evaluations, or data provided by the program.

Additional Considerations

- The general equation does not include any differential weights on dollar gains for participants based on income or other factors. The Robin Hood Foundation experimented with applying weights on earnings, but found no basis for choosing the weights. Additionally, their weighting methods did not affect the relative ranking of grant proposals (Weinstein and Bradburd, 2013).
- The general equation does not include the value of any cash or noncash public transfer payments, or benefits received or potentially lost due to increases in income. The only exception to this occurs when the program plays an active role increasing access to these benefits that otherwise would have been missed by participants, in which case we determine the amount of benefits to accrue to the program using Constellation's referral factor as described in section 6 in this document.
- To determine the counterfactual state of earnings or other outcomes in the absence of a program, we use average income and rates of outcomes for the target population in the Twin Cities area based on Census data. This average rate serves as an ad-hoc threshold for program impact.

2.3. Origins of the Benefit-Cost Model

Constellation's benefit-cost framework is inspired by two original sources. The majority of the methods and formulas presented in this document follow the benefit-cost model and technical documentation set forth by the Washington State Institute of Public Policy (WSIPP) as closely as possible. . We make frequent reference to WSIPP's documentation throughout this document. To our knowledge, the benefit-cost model developed by WSIPP is the first and largely the most rigorous application of benefit-cost analysis using meta-analytic evidence. However, Constellation's model differs from WSIPP's in that its BCRs are estimated only from the private participant's perspective, thus the results do not include any benefits accrued by the Government or other members of society. In addition, Constellation's model only counts the relevant operational costs of programming and operations, omitting public costs and other indirect benefits or costs caused by externalities associated with the program's activities. At this time, Constellation does not generate stochastic ranges for its BCRs.

The second source of inspiration is the metrics developed by the Robin Hood Foundation to estimate benefits of grantees. The models from WSIPP and Robin Hood are similar in that both rely on existing research evidence to estimate the impact of an intervention and then combine this evidence with local counterfactual states of the outcomes of interest to reach conclusions that are geographically relevant for decision making. In addition, WSIPP's model presents results that are applicable to the Washington State population while Robin Hood's and Constellation's BCRs refer to the impact of grantees at the municipality level (New York City and the Twin Cities Metro Area respectively).

2.4. General References

Weinstein, M. & Bradburd, R. (2013). The Robin Hood rules for smart giving. New York: Columbia University Business School Press. Retrieved from
http://cup.columbia.edu/book/the-robin-hood-rules-for-smart-giving/9780231158367

3. Calculating Per-Participant Impact

As discussed in the general equation in section 2.1, Q is the average per-participant impact of a program on an outcome of interest expressed in percentage points or percent change. Constellation estimates Q using a variety of different techniques, depending on both the evidence available and the type of outcome. In most situations, we obtain Q from:

- An evaluation of the program being assessed that shows the impact of the program on the outcomes of interest.
- Outcome data from the program.
- A suitable counterfactual base rate.
- Sufficient evidence from the research literature that the program is effective.

Formally, the impact Q is a function of an **effect size** and a **base rate**.

$$Q = F(ES, Base)$$

An effect size (ES) is a measure of how effective an intervention is. A base rate is the rate at which an outcome occurs absent intervention—for instance, the average high school graduation rate of a population, or the average depression score among depressed patients who are not treated. This equation is discussed in detail in WSIPP section 3.1 (see in particular equation 3.1.1).

3.1. Calculating Q Using a Quantitative Impact Evaluation of the Program

Constellation may be able to use an existing quantitative evaluation of a program to determine its impact and the corresponding economic benefits. The main requirements in these cases are:

1. The study of the program is of high quality and conforms to the requirements summarized in section 4. Evidence.
2. The study context and conditions under which the evaluation was conducted are still valid and applicable to the current situation of the program.

3.2. Calculating Q Using Outcome Data from the Program

In some cases, programs have outcome data that they collected and prepared internally or through the use of third-party evaluators. This outcome data may be used to estimate economic benefits for Constellation's benefit-cost analysis instead of (or in combination with) the impact evidence included in the metrics. In this section, we summarize the criteria to guide the choice of program-specific outcome data versus evidence from the literature when both types of information are available.

In general, we prefer to use program-specific outcome data over evidence from the literature. However, for Constellation to use program data, the information must satisfy the following criteria:

- The evaluation methods used to generate the evidence must include the use of a reasonable comparison group or counterfactual.

- If the program's data consists only of outcomes on participants, (i.e. the reported data does not include a comparison group), Constellation must be able to find a suitable counterfactual to estimate the net impact of the program. In this scenario, Constellation would judge whether using the literature evidence is more appropriate than using the program's data combined with an external counterfactual.

- If a "pre-post" method is used, it must be reasonable to assume that pretreatment data is a good counterfactual.

- The research takes into account potential biases affecting the data generation or the outcomes. These biases may include self-selection into the program, program selection ("creaming"), omitted variables, endogeneity, or spuriousness in the association of dependent and independent variables.

- The measured outcomes must be monetizable using Constellation's standard benefit-cost analysis. In most cases, this means that there is a way to logically connect the outcome to improvements in income, health, or reduced expenses.

- The evidence must be valid for the current context in which the program operates.

In some cases, the program data can simply replace the effect size component in an existing metric. We must pay attention to the units in which the outcome is expressed in the data and ensure it matches the metrics' original units. We will do our best to convert the program data to units that can be used in the benefit-cost analysis.

3.3. Calculating Q Using Existing Research Evidence

The key assumption when estimating Q by using existing quantitative research is that the program being evaluated and the programs assessed in the literature must be reasonably comparable. In other words, we assume that the grantee would have an effect on outcomes similar to the results in the reviewed studies. Constellation determines the appropriateness of the evidence based on several factors, including the similarities between the evaluee's program and researched program's design, the treated population, environmental conditions, and the scope of the outcomes measured. The detailed description of how Constellation assesses the good fit of the evidence is summarized in section 4. Strength of Evidence.

In the function to estimate Q, the estimated ES combines a set of quantitative research findings about a single topic into a single standardized statistic. Since all evidence is now expressed in the same standard units, we can draw conclusions from studies measuring the same outcome in different ways. We use standard meta-analysis techniques to estimate effect sizes (see for example: Lipsey & Wilson, 2000). In many situations, the usable evidence consists of only one study. In such cases, we would determine whether the results of the study should be standardized, or if they could be used directly or with some mathematical transformation to allow for adapting the paper's result to the local counterfactual rates.

We use two different estimation forms to compute Q = F(ES, Base)—one for dichotomous outcomes, such as graduation rates, and one for outcomes that are continuous in nature, such as test scores.

Calculating Q for Dichotomous Outcomes

For dichotomous outcomes, such as high school graduation or disease status measured as a binary present/not present, we use the equation:

$$Q = \frac{e^{ES*1.65} \times Base\%}{1 - Base\% + e^{ES*1.65} \times Base\%} - Base\%$$

where ES is the Cox effect size, and Base% is the rate of the outcome without intervention, as a percent. The Cox effect size is $ES_{cox} = \frac{ln(OR)}{1.65} = \frac{ln[\frac{P_t(1-P_c)}{P_c(1-P_t)}]}{1.5}$, where OR is the Odds Ratio, P_c is the percent of the control group with the outcome, and P_t is the percent of the treatment group with the outcome. When only one study is used as evidence and it reports results as odds ratios, we can instead use the simplified equation: $Q = \frac{OR \times Base\%}{1 - Base\% + OR \times Base\%} - Base\%$).

For further detail, see WSIPP section 3.2b.

Calculating Q for Continuous Outcomes

When the outcome is continuous—for example, test scores or levels of health represented in a scale—we estimate Q using:

Q = ES * Base Rate

where ES is the standardized effect size such as Cohen's d, or Hedge's g, and the Base Rate is measured in standard deviations. For further detail, see WSIPP section 3.2a.

In some cases, we rely on a single study to estimate Q. We use the estimation formulas for Q following the procedures described above. However, if the base rate for the population being served by the program is reasonably similar to the counterfactual rate in the study, we may use the result from the study directly, as long as we are able to estimate the effect in terms of percentage point increase. For example, if the study shows the intervention generates an 8% increase in the outcome for a population with similar levels of the outcome in absence of the program and comparable demographics, we can simply use the 8% as our Q value.

3.4. Linked Outcomes

In some metrics, the final economic benefits of a treatment are obtained indirectly by linking the initial impact of the treatment on an intermediate outcome to the subsequent impact of the intermediate outcome on a final monetizable outcome. For example,

Treatment → Outcome1 → Outcome2 → Economic impact

In these cases, we simply calculate multiple Q values—one for the impact of the program on the intermediate outcome, and one for the impact of the intermediate outcome on the final outcome—as described above and multiply the two together. Usually, we do this via two separate components within a metric, so that all calculations are visible and explicit. Occasionally, however, we may combine all calculations into a single Q value. For further discussion, see [WSIPP 3.3, 3.4, and 3.5].

3.5. Counterfactuals

Constellation values the benefits that accrue *solely due to program participation*. This means considering the counterfactual, or what would have happened absent program participation. Often this is accomplished by using the appropriate base rate, as discussed above. In this section we summarize some situations where determining the appropriate base rate is difficult and we need to find alternative measures for a counterfactual.

One situation worth considering is when services are so prevalent that their impact *can't* be removed from the base rate, due to lack of data. Consider again the example of a tutoring program aiming to increase high school graduation. To apply the appropriate counterfactual, we would want to subtract from the program's graduation rate the graduation rate of those students with an absence of tutoring. To do this we use the general graduation rate for low-income students in the Twin Cities. However, with various student support programs in place across the metro, it is likely that this low-income graduation rate already includes some of the impact of the program we're evaluating and comparable services, so the true counterfactual may be lower than what we observe. In this case, depending on the data available, we may be able to adjust the counterfactual slightly to account for this, or we may have to use the general low-income graduation rate as our counterfactual, and accept that it will be slightly higher than the "true" counterfactual.

4. Evidence Selection

Constellation uses the best available evidence to create its metrics. The general process we follow is:

1) Determine the scope, characteristics and target population of the interventions implemented by the evaluee;

2) Gather research evaluations that have been done on each intervention; and

3) Use meta-analysis or other quantitative techniques to draw an overall conclusion about the average effectiveness of an intervention.

This evidence is then combined with local demographics and the economic value of the outcomes assessed to produce a metric for the per-participant benefit of each intervention.

However, with evaluees doing work in unique and innovative ways that may not have been robustly studied, there can be substantial variation in the quality of the "best" evidence across different areas. In general, Constellation follows these practices:

- Constellation conducts primary searches to locate studies using a Research Protocol. See section 4.5.

- Constellation prefers to use meta-analyses that combine results from many different studies. WSIPP is the preferred source.

- If meta-analyses are not available, Constellation combines the results of multiple studies and conducts its own meta-analysis estimations.

- Constellation prefers data from randomized control trials (RCTs) or high-quality quasi-experimental design studies (QEDs). When no published research is available, Constellation may also consider evidence from other sources, such as private reports, if they are sufficiently high-quality.

- We examine studies from peer-reviewed and non-peer reviewed publishers as long as they meet our selection criteria.

After assembling the best available evidence, Constellation rates it as good, sufficient, or unusable in four areas: specificity of evidence, chain of logic, availability of evidence, and appropriateness of effect or measure. When Constellation assigns a sufficient rating, it also includes a comment noting the limitations of the evidence.

4.1. Specificity of Evidence

Constellation considers how well matched the interventions or programs studied in the research are to the intervention or program for which we are writing a metric. An example of an ideal match would be an organization that delivers a specific home-visiting curriculum with fidelity to the model accompanied by research evidence studying an implementation of that exact curriculum. A less ideal match would be if no research was available for the specific home-visiting curriculum that an organization implements, but if there were studies of home visiting programs with roughly similar goals and dosage. If the only available evidence was a meta-analysis that looked at home visiting programs with greatly varying goals and dosages, we would likely write a generic home-visiting metric.

4.2. Chain of Logic

This criterion looks at how directly or indirectly the research can draw a line from the intervention to the monetized benefits. This considers the number of "links" between intervention and benefits, but also the theoretical and empirical strength of those links. For instance, if an RCT of a home visiting program shows increased high school graduation rates, Constellation can link the intervention to increased lifetime earnings with few links in the chain. An example of a less direct link would be if one study shows that an after-school program is shown to reduce disruptive classroom behavior and another study shows that reducing disruptive classroom behaviors increases high school graduation. The most important factor when determining the strength of these linked effects is the quality of the evidence. For example, we would consider linked outcomes based on RCTs as strong links, and we would rate any benefits estimated using these strong links as less uncertain than benefits based on correlational studies.

4.3. Availability of Evidence

This criterion considers both the quantity and quality of available evidence. Our preference is for a published meta-analysis, or for multiple RCTs or high-quality QEDs that Constellation can combine into a meta-analysis. Acceptable but less ideal evidence includes choosing a single study because other available studies report their results in a way that cannot be combined into a meta-analysis, or opting for correlational evidence in cases where absolutely no causal evidence exists and there is a compelling theoretical basis for assuming causation.

4.4. Appropriateness of Effect or Measure

This criterion rates the extent to which the data available from the research is a match to what the metric needs. Ideally, the effect or statistic will be exactly what's called for, for example, the impact of a program on high school graduation for a metric valuing increased graduation rate. Also acceptable are approximations that require only reasonable assumptions, for instance, using the incidence of HIV among people who inject drugs nationally when the metric calls for that incidence among the local population of people who inject drugs. When nothing better is available, we may use less ideal approximations, such as considering a broad meta-analytic category of "educational success" as a proxy for graduation rate. A weaker match may also occur when the evidence refers to levels of an outcome whereas the final outcome to be monetized may be the prevalence of the outcome. For example, the evidence regarding the effectiveness of a diabetes treatment may be expressed as A1C levels while the local counterfactual data is the prevalence in the targeted population. In cases like this, Constellation would need to make stronger assumptions to use these results.

4.5. Research Protocol

The Research Protocol is used by the Research Associate and any other member of the Metrics and Evaluation team who is seeking out evidence to support a developing metric. The protocol may also be used when reviewing or updating an existing metric.

Searching for Evidence

First, the topic is defined for the area of research, desired outcomes, and research terms. For example, a metric on treatment for depression may include search terms such as "mental health," "therapy," or "antidepressants," in tandem with obvious terms like "depression."

Resources checked first include:

Mendeley: Constellation's database for storing research papers previously identified by the Metrics and Evaluation team as high quality and either used in current metrics or saved for future use.

WSIPP: Washington State Institute for Public Policy's library of benefit-cost analysis outcomes covering a wide range of topics from workforce development to health care. To navigate, hover over the "Benefit-Cost" tab and click the "Results" option in the drop-down menu. To understand their method of calculating effect sizes, hover over the "Benefit-Cost" tab and click the "Technical Documentation (PDF)" option.

Journal of Benefit-Cost Analysis: An interdisciplinary and wide-ranging journal exclusively focused on benefit-cost analysis.

Google Scholar, "Additional Resources" (see appendix), or other databases may be used to conduct a specific search on the topic of interest using the terms previously specified.

Terms related to benefit-cost analysis are included. For example: cost-effective, cost-benefit, QALYs, cost-efficient, ICER (Incremental Cost Effectiveness Ratio), and iteratives of these terms.

Randomized control trials (RCTs) and meta-analyses are prioritized.

Saving Sources

At the end of each funding round sources are uploaded to Mendeley and tagged by metric name, area of focus, and other identifying topics.

5. Procedures for Considering the Percent of Participants Under the Poverty Threshold

Constellation's BCR must reflect benefits accrued by low-income participants compared to the relevant cost to generate these benefits. Some organizations serve individuals of various income levels, including some that would not be considered low-income earners by Federal guidelines. Thus, a Constellation grant could potentially be used to finance services for higher income families and individuals. A BCR that does not account for this issue would confound the impact of the grant and include returns to the higher income individuals or mix the costs of serving these two groups.

Whenever we suspect that a fraction of an organization's participants may not be categorized as low-income, we must find the percent of participants defined as low-income to accurately measure the benefits accrued to them and the respective cost of providing services to this group. In general, we prefer to determine this information using the organization's own criteria to define and select low-income participants. This is our preferred method since the Federal guidelines may be an imperfect approximation to define poverty. We give more weight to the organization's expertise in determining the needs of the community and identifying the relevant target population. However, in the absence of a formal definition from the organization, we use the Federal guidelines to estimate this information. The specific steps we follow to determine the number of low-income participants is summarized below.

5.1. Considering the Proportion of Low-Income Participants Served By an Organization to Determine the Merit of an Application

As part of an organization's initial application, we request their best estimate of the percentage and number of low-income participants as defined by the organization and/or using Constellation's definition based on Federal guidelines. Constellation defines "low-income participants" as: Individuals or families with income under 200% of the federal poverty line (FPL). However, this threshold must be used with caution as noted above.

See table <FPL Table below>.

We use the resulting number of low-income participants and the rest of the information from the application to determine the merit of the application. We do not require a minimum percentage or number of low-income individuals served. However, an organization seeking funding from Constellation must show that it has a relatively high impact on low income individuals.

Explanation: It is important to have an idea of the actual number of low income participants because the raw percentage may be misleading for our purposes. For example, in a large organization serving 10,000 participants, only 10% may be low-income participants, which is 1,000 individuals. This is more than what many smaller organizations serve. But, even with smaller numbers than the ones in this example, we may find the intervention worth evaluating for other reasons, including the novelty of a model, the intrinsic value of serving a particular vulnerable group, closeness of the intervention to one of our "Focus Areas," etc. So, a pre-established threshold is not advisable. Instead, we need to look at the whole application to determine the evaluability of an intervention.

5.2. Considering the Proportion of Low-Income Participants Served by an Organization to Adjust Benefits

If the organization is selected for a full evaluation, we need to determine as accurately as possible the percentage and number of low-income clients by each program, activity, or metric.

If the organization does not have a clear definition or data to estimate a precise number or percentage of low-income participants, IOs can use the FPL to determine this number. However, this threshold must be used with caution as noted above.

1. If the percentage or number of low-income participants is known by metric or program activity, use that number in each metric.
2. If the specific numbers of activity/metric are not known or hard to estimate, you may use the percent of total individuals defined as low-income participants and apply this to the number of participants in each metric or adjust the total benefits (the sum from all metrics) using the percentage of low-income participants. IOs and the M&E team should determine which approach is a better representation of the organization's impact.

Note: If the percentage of higher-income participants is very low, the IO may deem this level negligible for the purpose of the BCR estimation. IOs must consider how sensitive the final BCR could be to excluding even a small number of higher-income participants.

5.3. Considering the Proportion of Low-Income Participants Served by an Organization to Adjust Costs

The main question to guide this part of the process is: what is the cost of generating the benefits accrued by low-income participants?

1. Determine if it is possible and feasible to estimate the specific costs of serving low-income participants. This estimate must be based on actual accounting and programmatic data and be as free of assumptions as possible.
   a. If so, use the low-income-specific cost estimates in the BCR. Note: the relevant overhead costs must be estimated and added to the total cost.
   b. If not, go to step 2.
2. If costs of services are difficult to disaggregate across income levels, determine if it is reasonable to assume that the average cost per participant is the same across income levels, i.e., the organization provides the exact same service to all participants.
   a. If so, estimate the average cost per participant and multiply by the unduplicated number of low-income participants. Note: the relevant overhead costs must be estimated and added to the total cost.
   b. If not, go to step 3.
3. In some cases, it may be expected that the average cost per low-income participant is different from the average cost of serving the higher income participants. For example, a low-income participant may come with more unmet needs that require more care or services. If this is the case and there is no accounting or programmatic data to determine the cost of serving low-income participants, we work with the organization to establish a subjective factor that reflects how much more work or resources a low-income participant requires over a

higher-income participant. This factor is a percentage over the average cost per participant. For example, if the average cost per participant is $100, and the program leaders estimate that it takes 25% more work to serve a low-income participant, we would adjust the cost by multiplying it by (1 + 25%). Using this equation, the adjusted cost per low-income participant would be: $100 x 1.25 = $125. Relevant overhead costs must be estimated and added to the total cost. Note that the total cost determined using this procedure should not exceed the total cost of the organization, which may become an issue if the adjustment factor or the percentage of low-income participants are too high.

Table 1: Annual Income Limits[1]

Household Size	Annual Income Federal Poverty Guideline (FPG)	125% FPG	150% FPG	175% FPG	200% FPG	250% FPG	300% FPG	350% FPG
1	$14,580	$18,225	$21,870	$25,515	$29,160	$36,450	$43,740	$51,030
2	$19,720	$24,650	$29,580	$34,510	$39,440	$49,300	$59,160	$69,020
3	$24,860	$31,075	$37,290	$43,505	$49,720	$62,150	$74,580	$87,010
4	$30,000	$37,500	$45,000	$52,500	$60,000	$75,000	$90,000	$105,000
5	$35,140	$43,925	$52,710	$61,495	$70,280	$87,850	$105,420	$122,990
6	$40,280	$50,350	$60,420	$70,490	$80,560	$100,700	$120,840	$140,980
7	$45,420	$56,775	$68,130	$79,485	$90,840	$113,550	$136,260	$158,970
8	$50,560	$63,200	$75,840	$88,480	$101,120	$126,400	$151,680	$176,960
9	$55,700	$69,625	$83,550	$97,475	$111,400	$139,250	$167,100	$194,950
10	$60,840	$76,050	$91,260	$106,470	$121,680	$152,100	$182,520	$212,940

[1] "Office of the Secretary Annual Update of the HHS Poverty Guidelines." Federal Register 88, no. 12 (2023): 3424-3425.

6. Procedures to Estimate Benefits from Third-Party Providers

Many poverty-fighting organizations assessed by Constellation refer participants to third-party providers for services that contribute to the mission of the organization. For example, a supportive housing program may refer participants to mental health providers or job-assistance programs. To obtain a fair assessment of the total impact of the applicant, Constellation needs to account for the share of benefits generated through referred services that can be attributed to the organization's operations. We define this portion of benefits using a referral factor. This factor is applied to every metric associated with referred services.

We use the following steps to determine whether to include benefits from third-party services and what fraction should be added to the BCR calculations:

1. Identify the outcomes associated with the organization that are generated by third-party providers. This conversation starts during the application process and the final list of third-party services is confirmed through site visits and the data collection process.
2. Determine the role of the organization in connecting participants to third-party providers. This is confirmed during site visits.
3. Determine if the third-party service is well captured by our metrics to allow us to apply our framework.
4. Determine the number of participants being referred to and the percentage of those actually receiving third-party services. This information is gathered during the data collection process.
5. Determine the referral factor and apply it to the relevant metrics. For example, if a housing program refers participants to a mental health provider for care, with total health benefits estimated at $10,000, and Constellation establishes a third-party adjustment factor at 50% for this service, then $5,000 is attributed to the housing program and used in the BCR analysis as coming from their mental health referrals.

6.1. Referral Factor for Third-Party Services

We use the following table to determine the referral factor. The table shows the referral factor associated with the level of involvement of the organization in increasing access to third-party services and how much information on the referrals is available. For example, an organization may be weakly involved in a referral when it only provides information to its participants about the availability of the service but does not provide any further support to ensure participants receive the service. The referral factor also depends on the availability of data regarding how many participants actually receive the referred services. The referral factor ranges from zero to a maximum of 50%. We summarize the logic and evidence we use to determine these values in the following sections.

Table 2 : Adjustment Factor to Estimate Benefits from Referrals to Third-Party Providers

Adjustment factor for third-party services	Role of referring organization

			Weakly involved	Somewhat involved	Highly involved
			10%	25%	50%
Availability of data on referrals	Third-party information is made available, but records are not available.	Probability of receiving service: 0%	0%	N/A	N/A
	Only the number of participants who are referred is known.	34% to 52%	3%	9% - 13%	N/A
	Number of participants who receive service is known.	75%	N/A	N/A	38%

6.2. Methods to Determine the Adjustment Factor

The adjustment factor represents how much of the benefit from a third-party service is included in the evaluee's BCR. The referral factor is summarized as:

$$Referral\ Factor = (Prob.\ of\ Receiving\ Service\%)\ x\ (Attribution\%)$$

The referral factor combines the probability that a participant receives the third-party service and the amount of benefit attributed to the referring organization.

Probability of Receiving the Third-Party Service

We use the availability of information on referrals to determine this probability and use research findings from Boyum, et al., (2016) to inform our classification. Boyum, et al., (2016) shows that between 34% and 52% of individuals who receive information about a service through a referral do access the service provider. We developed three possible scenarios for this analysis as described in Table Z. In the first scenario, there is not enough data to determine the number of participants who received the referral. The most common scenario occurs when the organization knows how many participants were referred but does not know the number of participants who actually received the service. In a more ideal scenario, the organization has data on how many participants received the service after being referred.

Table 3: Probability of receiving services

Availability of data on referrals	Description	Probability of receiving service
The referring organization does not keep any records about who receives the referrals.	An example of this case is when information about third-party services is made available to all participants but data is not tracked. This is the baseline we assume as the counterfactual.	0%
The referring organization only keeps records about who receives the referrals.	This must be more than just making information available. In this case, Constellation estimates the probability of receiving the service is between 34% and 52%, depending on whether the referring organization makes the referral with knowledge about participants' eligibility to receive the service (Boyum et al., 2016).	34% and 52%
The referring organization tracks and records the number of participants who actually receive third-party services.	In such situations, we assume the middle point between 100% and 52%, or approximately 75%, as the probability of receiving the service (Boyum et al., 2016). In cases where we are particularly certain of the number of participants who actually received services, we may use 100%. This is determined on a case-by-case basis.	75%

Attribution of Benefits from Referrals

The attribution of benefits from a referral is intended to approximate the extent to which the referring organization can be thought of as being "responsible" for the participant receiving services. To our knowledge, there is no research evidence to objectively construct an attribution factor. Thus, Constellation determines this value subjectively using the following criteria.

This factor must always be significantly lower than 100%, even in an extreme case where an organization serves as the only liaison between the community and the third-party service providers. Though they may be increasing the probability of receiving services from zero to 100%, they are still not incurring the costs associated with actually providing those services. Constellation assigns a maximum contribution factor of 50% when the organization plays a very important role in increasing access to the referred service. The third column in Table 2 lists the amount of benefit we attribute to the referring organization depending on three levels of involvement in the referral.

We assign contributions of 10%, 25%, and 50% for weak, somewhat, and highly involved organizations, respectively. These percentages are subjectively defined by Constellation and may be modified by impact officers on a case-by-case basis.

Table 4: Level of Involvement in Referrals

Level of involvement in referral	Description	Contribution to final benefit
Weakly Involved	Information about third-party providers is provided, but no formal referral is made.	10%
Somewhat Involved	Participants are provided some encouragement or education to motivate participation. The referral service is optional and not highly encouraged.	25%
Highly Involved	Participants are unlikely to achieve outcomes from the third-party provider but for the referring organization. The referring organization provides encouragement or education to motivate participation. The referring organization likely provides some help in physically accessing services, such as providing them on-site or providing transportation to the service. There may be a memorandum of understanding or a contract in place to guide and document referral procedures. The program's core mission is to help individuals access social services that they otherwise may not have received.	50%

Determining the Adjustment Factor

The final step to determine the adjustment factor is to combine the data on the probability of receiving the third-party service and the level of contribution of the referring organization to the final benefit. Table 2 summarizes the final adjustment factor for each level of participation of the referring organization and the available information on the referrals. The inner cells in Table 2 show the proposed adjustment factor for each scenario.

For weakly involved organizations, the only possible option where any benefit is counted toward the BCR is when only the number of referrals is known. In this case, the adjustment factor is 3%, which means that after computing benefits for the intended outcome, only 3% of the benefit is counted towards the BCR of the referring organization. This factor results from multiplying the probability of receiving services given the availability of data by the level of involvement of the referring organization (34% x 10% = 3%).

For organizations that are somewhat involved in the referral process, the only possible option where any benefit is counted is only when the number of referrals is known. In this case, the adjustment factor can be estimated at 9% or 13% as shown above. IOs choose one of these values depending on whether the referring organization knows and uses eligibility information when making the referrals, which leads to a higher adjustment factor.

For organizations that are highly involved in the referral process, the only possible option where any benefit is counted is when there is data about the number of participants receiving the third-party service. We assume that organizations are not highly involved in the referral process if they don't actively track service utilization. In this case, the adjustment factor can be estimated by multiplying the actual percent of referred participants receiving services by the percent contribution of the benefit (50% for that level of involvement). When the actual fraction of participants receiving services cannot be computed, we assume a 75% chance of receiving services with the resulting adjustment factor of 38% (75% x 50%).

A Note on Overestimation of Total Societal Benefits

Many of the programs Constellation evaluates refer participants to other providers. Furthermore, it is also possible that a fraction of an evaluee's participants are referred from other organizations that Constellation may have evaluated. Thus, from a societal perspective, there is a small chance of overcounting benefits when evaluating organizations within an ecosystem of referrals as long as there is not crowding out of participants from the additional number served from referrals. Crowding out may occur if third-party providers are working at capacity and if serving referrals changes the number of persons who get service and thus societal benefits may be diminished. However, while Constellation estimates the BCR from the perspective of the individual participants, we do not count all societal benefits and costs. So, while two estimated BCRs may contain some overlapping benefits, we interpret each BCR as an independent indicator or the impact of a program. The potential overestimations are a small fraction in most cases.

7. Public Transfers Administered by Nonprofits

Many nonprofits serve as either administrators or direct providers of public benefit programs, such as rental assistance or other cash-type transfers. In most cases, we apply a referral factor to the amount of public benefits administered by the organization as described in a previous section. More precisely, when determining the referral factor, we focus on answering to what extent the organization increases the value of benefits received by low-income Twin Cities residents. If an organization increases take up of a public entitlement, the organization may get full credit for that benefit. On the other hand, if that person would have gotten the benefit anyway through another channel, then assigning no benefit is appropriate.

7.1. Rental Assistance

In general, we don't include any pass-through rental assistance dollars as benefits in the BCR estimation. For organizations providing pass-through rental assistance, we still need to adjust the costs estimates to compute the BCR. Since organizations are incurring some costs to deliver these transfers, we must reduce the cost used in the BCR. In these cases, we use the following formula:

$$Cost\ Deduction = \ (Number\ of\ individuals\ receiving\ rental\ assistance) \times (Administrative\ Costs)$$

Whenever possible, we use program cost data to determine this overhead cost. In cases where this information is not available, for example, if the organization does not know how many participants receive the assistance and only knows the total cost, the participant number is estimated as:

$$Number\ of\ participants = \frac{(\$\,Total\ value\ of\ assistance)}{(\$\ average\ assistance\ per\ person)}$$

If the average assistance per person per year is not known, use $6,000, which is the average amount received based on all public rental assistance programs in MN.

8. Procedures to Estimate Income

8.1. General Assumptions for Annual Income Estimates

- These calculations are based on the 5 Year ACS PUMS Data (2019) (American Community Survey - Public Use Microdata Sample)
- All estimates refer to the Twin Cities Metropolitan Area
- The earnings include wages and self-employment income
- All the earning values are rounded up to the nearest dollar and adjusted to 2021 U.S. Dollars using the average CPI for all urban consumers as available through this link: https://data.bls.gov/timeseries/CUUR0000SA0
- The earning figures are adjusted for state and federal taxes
- The lifetime earnings are calculated using an annual discount rate of 3%
- Missing data were imputed using the mean value imputation method (this was done only for immigrant samples)
- Earners denote individuals who earned $1 or more
- Non-earners denote individuals who earned $0 or less
- Non-immigrant denotes individuals born in the United States
- Immigrant denotes individuals born outside the United States
- Low-income denotes individuals whose annual earnings are below 1.85 times the FPL guidelines based on the number of people in the household as mentioned here: https://aspe.hhs.gov/topics/poverty-economic-mobility/poverty-guidelines/prior-hhs-poverty-guidelines-federal-register-references/2021-povert

8.2. Procedures to Estimate Lifetime Earnings by Educational Attainment

Our methods for estimating the impact of educational attainment on lifetime earnings are a simplified version of the model developed by the WSIPP. Technical documentation for the WSIPP approach can be found in WSIPP section 4.2b and section 4.9a.

The general procedure consists of estimating the average income by educational attainment for the populations of interest. Since we use microdata in a cross-section format, we first obtain individual level data points by age, with which we can produce a smooth series by each educational attainment using predicting values. We then adjust the resulting lifetime series by the probability of death by age, deduct taxes, and discount to present value. We update these calculations every 3 years, or whenever the Census reports a newer version of the data.

Sources

We derive all earnings-related estimates from the U.S. Census Bureau's American Community Survey (ACS), 2019; 5-year estimates using the Public Use Microdata Sample (PUMS), which provides cross sectional data for earnings by age; educational status; and other characteristics. The variables used and detailed in the estimation protocols are available upon request. The sample is restricted to persons ages 18 to 65 inclusive, the 7-county Twin Cities metro area, and weighted by the ACS person weight. We use tax incidence rates from the U.S. Department of the Treasury by income decile and the Minnesota Department of Revenues. We adjust income levels using life expectancy data from the Center for Diseases and Control (CDC).

Earnings by Educational Attainment

We use the ACS variable for educational attainment by the highest level completed, to subset the sample by education. We perform the calculations described above using subsets of the data sample for four educational status groupings (and two subset groupings):

- Those who completed lower than 8th grade
- Those who completed at least 8th grade but did not report completing high school
- Those who reported completing high school with a diploma
- Those who reported completing high school equivalence
- Those who reported completing high school, regular or equivalence
- Those with some college but no degree
- Those with an Associate's degree
- Those with a Bachelor's degree

For each of these groups, we replicate the regressions and modeling to determine separate earnings by age distributions. The average income reported is for all people at each age, not just for those with earnings. Thus, the ACS data series we include in the model measures both earnings of the earners and the rate of labor force participation.

Table 5. Lifetime Income by Educational Attainment - Twin Cities Metro Area - General Population (2021 US Dollars - After Tax Adjustment)

Educational Level	General Population
Masters, Professional, or PhD degree	$2,266,155
Bachelor's or Higher*	$2,025,872
Bachelor's degree	$1,888,045
Associate's degree	$1,496,727
Some college	$1,296,243
High School**	$1,195,057
No High School***	$767,791

* Bachelor's or Higher includes those with Bachelor's degree, and those with Masters, Professional, or PhD degree

** High School includes those with Regular high school diploma and those with GED diploma

*** No High School includes those with less than High School but more than 8th grade and those with less than 8th grade

8.3. Procedures to Estimate Annual Income for Selected Populations and Educational Attainment

We use average annual income to evaluate many program's outcomes, for example, increases in employment, health, housing, or workforce programs. Our annual income estimates also serve as proxies for counterfactual states of these outcomes, thus we estimate this statistic for several populations of interest and demographics. In general, we follow the same procedures described in the lifetime section to estimate annual averages. The difference is that we do not estimate predicted lifetime series, but rather simple averages. Specific characteristics of each estimate are listed in the metric's details. Table X shows the most commonly used income estimates we use. In addition, we estimate annual income for earners or non-earners whenever we want to make explicit the employment rates in the metric, like immigrant workers or formerly incarcerated individuals. In some cases, we estimate ad-hoc income levels for subgroups that are relevant to a particular metric or program such as individuals experiencing homelessness, or by gender or other characteristics.

Table 6. Commonly Used Annual Income Averages

	All levels of education		Less than High School		Only High School	
	General pop	Low-income	General pop	Low-income	General pop	Low-income
Ages 18-65	$70,648	$24,195	$33,894	$22,767	$52,979	$28,061
Ages 18-25	$26,447	$19,354	$15,122	$12,579	$22,602	$19,859
Ages over 50	$78,373	$18,476	$39,726	$20,634	$62,349	$24,522

9. Education Interventions, Methods, and Assumptions

The main source of benefits generated by educational interventions come from increased probability of earning a degree; the intervention can be linked to this outcome through academic and non-academic activities. For example, we link an increase in standardized test scores as well as social emotional learnings' impact on behavior to probability of HS graduation. Constellation determines the dollar amount to attribute to an educational intervention by taking into account two factors: 1) the net impact of the intervention on the probability of achieving the educational outcome, and 2) the adjusted monetary value of that outcome. To determine the effect of the organization on the outcome of interest, we require a suitable counterfactual. Generally, we use the graduation rate of the population of interest for the relevant attainment level as the counterfactual. All counterfactuals used are reported in the respective educational metrics as well as the monetary value associated with these outcomes. We use computed income estimates following the procedures described in Section 8.2.

We also incorporate a causation factor to adjust the net income gains associated with each level of educational attainment. This factor accounts for the fact that not all the differences in earnings observed across levels of attainment are attributable to the completion of the degree. Since Constellation's earnings regressions do not include demographics or intervention variables, the adjustment factor helps to approximate the net impact of the degree. We use the causation factors noted by WSIPP in exhibit 4.8.5.

Many educational programs collect graduation rates for their participants. We use the procedures described in previous sections to determine whether to use program data or effect sizes from the research literature.

10. Supportive Housing Methods & Assumptions

Constellation's housing metrics have a number of moving parts that make them unique within our framework of metrics. In this section, we summarize the assumptions and methods used to construct housing metrics. The components of housing metrics we account for include the duration of program and benefits, participant numbers, and counterfactuals.

10.1 Two Types of Supportive Housing Outcomes

Broadly speaking, outcomes for supportive housing metrics can be "lifetime" outcomes—like high school graduation or individuals who have experienced child abuse—which accrue only once, or "single-year" outcomes—like an increase in annual income or avoiding chronic health issues while housed—which can accrue multiple times. These outcomes determine how duration factors and counterfactuals are approached.

10.2 Housing Programs with Lifetime Outcomes

For housing metrics with "lifetime" benefits, the outcome can only be achieved once, so if a program's typical duration is more than one year, the benefit should be divided by the average duration of the program to reflect the fact that some of the costs contributing to that outcome are accrued in a different program year.

10.3 Housing Programs with Single-Year Outcomes

For metrics with "single-year" outcomes, a year's worth of benefits are ascribed, and the benefits are directly attributable to a participant being housed during that year. Therefore, the metric can be applied to the same participant, year-after-year, without double-dipping. This short-term duration of benefits is also observed in the majority of evidence used in the housing metrics.

10.4 Number of Participants

Since benefits are calculated as the amount of benefit accrued by a participant from a single year of housing, a participant who is served for only part of the program year would accrue less than the full benefit amount. To account for this, Constellation assigns benefits according to the number of participant years served.

- One "participant year" is a participant who receives services for one entire year.
- To determine the amount of participant years generated by a single participant, divide one by the duration in years of that participant's stay. For example, a single participant who is housed for six months equates to .5 participant years. A participant who is housed for nine months equates to .75 participant years.
- In cases where each housing unit can house exactly one occupant, the number of participant years is equal to the number of units (assuming the program operates at capacity all year), regardless of any turnover that occurs.

- In cases where one unit may house a variable number of participants (e.g., 1 or 2 adults; 1 or more children), for ease of calculation, the number of participant years should be the average of the number of participants housed between the beginning of the program year and the end of the program year. This approach accounts for the unit turnover during the program year and may result in different numbers of participants being served at different times.
- This approach is applied regardless of the type of outcome (lifetime or single year).

10.5 Estimated Impact of Supportive Housing

Most supportive housing metrics begin with a Q that determines the percent of participants who achieve the intermediate outcome—being housed—because of the program in question. This impact is the percent of participants housed minus the percent who would be housed anyway:

Q = (Actual number housed by program) - (Counterfactual)

Actual

The first part of the Q value is the percent of participants who achieve the outcome. For single-year metrics, this is 100%, since, by virtue of participating, participants are housed long enough to obtain the benefit. However, for lifetime metrics, the benefit requires that the participant remain housed for a much longer period of time. The percent of participants who achieve that lifetime outcome will be smaller than the total number of participants, reflecting the fact that some portion of participants will again become homeless after the intervention ends. See column 2 of table X.1.

Ideally, if a program tracks outcomes of participants after they exit supportive housing, Constellation uses these outcomes to determine what portion of participants retained stable housing. Programs (and literature, as discussed below) are unlikely to have outcomes data further out than one to two years after leaving housing. Absent better data, it is assumed that a person who remains stably housed for one to two years after leaving supportive housing remains stably housed in perpetuity.

Note that there may be exceptions. If a permanent supportive housing program has a long average duration and does not have a time limit on how long participants can stay, it may be reasonable to assume anyone exiting the program is exiting into stable housing. This is considered on a case-by-case basis.

Counterfactual

In all cases, the actual rate of the percent of participants who would have achieved the desired outcome (stable housing) even absent the intervention is subtracted. See column 3 in table X.1.

The numbers shown in table X.1 from Gubits and Rodriguez should be regarded with caution. Both Gubits and Rodriguez report a value for the comparison population (ranging from about 48% to 76%); however, these numbers are unintuitively high. In the case of Rodriguez, the comparison group is not a true control, it is people who enter emergency shelter and do not subsequently receive supportive housing, so this population may have been in far lesser need of supportive housing services than those who were served. In the case of Gubits, the treatment and control group refer to who was prioritized for services, not who received them. Some people in the treatment group did not receive any kind of supportive housing and some in the control group did. In both situations, the number of expected untreated participants who achieve stable housing is artificially inflated. Lacking better data, it is assumed this number is 50%, though in reality it may well be lower.

The counterfactual varies based on population served. For instance, in the case of youth who have run away from unsafe home situations, it may be reasonable to assume that 0% would be able to achieve safe, stable housing without some sort of intervention.

The considerations mentioned above should be used as starting points, but in cases where it is reasonable to believe either or both parts may be significantly different than the literature suggests—perhaps because of the particular population served or because of contexts that inform the likelihood of finding housing absent some intervention—it may be sensible to adjust these numbers.

Table X.1: Supportive Housing Q-values

	Actual (% of participants who achieve the desired outcome)	Counterfactual (% of participants who would have achieved the desired outcome even without intervention)	Q (Actual - Counterfactual)
Single-year outcomes (all housing types)	All populations: 100%	General population: 50%* Vulnerable or underserved populations: 0%**	General population: 50% Vulnerable or underserved populations: 100%
Lifetime outcomes	TH, allpop: 71.0% RR, allpop: 69.9% PSH, allpop: -***	All housing types, general population: 50%* All housing types, vulnerable or underserved populations: 0%**	TH, genpop: 21.0% RR, genpop: 19.9% PSH, genpop: 15.3% TH, vpop:71.0% RR, vpop: 69.9% PSH, vpop: 65.3%****

TH=Transitional Housing, RR=Rapid Rehousing, PSH=Permanent Supportive Housing, genpop=general population, vpop=vulnerable population, allpop=all populations (i.e., general or vulnerable)

*The percent of the general population that would achieve safe and stable housing even without intervention is difficult to discern. Both Gubits and Rodriguez report a value for the comparison population (ranging from about 48% to 76%); however, these numbers are unintuitively high. In the case of Rodriguez, the comparison group is not a true control, it is people who enter emergency shelter and do not subsequently receive supportive housing; this population may have been in far lesser need of supportive housing services than those who were served. In the case of Gubits, the treatment and control group refer to who was prioritized for services, not who received them—some people in the treatment group did not receive any kind of supportive housing, and some in the control group did. In both situations, we expect the number of untreated participants who achieve stable housing to be artificially inflated. Lacking better data, we assume this number to be 50%, though we recognize in reality it may well be lower.

**"Vulnerable or underserved populations" refers to any population which is very unlikely to achieve safe and stable housing without community support, e.g., runaway youth, youth who have been sexually exploited, etc.

***The evidence for permanent supportive housing does not report percentages stably housed for treated and untreated groups, rather it uses OLS regression to report on the impact of receiving supportive housing and finds a 15.3 percentage point difference. This essentially gives us the Q value while skipping over the actual and counterfactual rates.
****We assume that for vulnerable populations the percentage point difference increases by an additional 50 percentage points over the regression-based Q for supportive housing.

Supportive Housing References

Gubits, D., Shinn, M., Bell, S., Wood, M., Dastrup, S. R., Solari, C., ... & Spellman, B. (2015). Family options study: Short-term impacts of housing and services interventions for homeless families. US Department of Housing and Urban Development, Office of Policy Development and Research.

Rodriguez, J. M., & Eidelman, T. A. (2017). Homelessness Interventions in Georgia: Rapid Re-Housing, Transitional Housing, and the Likelihood of Returning to Shelter. Housing Policy Debate, 27(6), 825–842. https://doi.org/10.1080/10511482.2017.1313292

11. Health Interventions, Methods, & Assumptions

11.1. General Approach to Assess Health Interventions

Constellation's benefit-cost analysis includes benefits from health improvements. We consider the impact of treatments given by health providers or health benefits derived indirectly from other outcomes such as academic achievement or employment. In some cases, we assess health improvements associated with screening services that help with early detection and diagnosis of diseases. Benefits from screening are usually adjusted such that the future benefits from the actual treatment of the disease is accounted for.

11.2. Assumptions

- Constellation uses quality-adjusted life years (QALYs) as our preferred measure of economic value for health outcomes. The QALY is a measure of both the quality and the quantity of life lived. One QALY equates to one year living in perfect health. See section Y for more details on our use of QALYs.

- We do not include any adjustments for persistence, recurrence, or remission of mental illnesses in the estimation of future benefits. Rather, we assume that these are accounted for in the QALY estimates. Thus, in most cases, we limit benefits to the short-term. In estimates of lifetime benefits, we use evidence from the literature that has already accounted for these factors.

- We estimate the value of a life using $50,000 per year value, and a discount rate of 3%. We combine these parameters with mortality rates and life expectancy data for specific populations.

- The duration of health benefits included in the BCR analysis depends on whether the health improvements are achieved from lifetime? treatment, for example, a vaccine that requires further boosters or treatments that last more than one year or substance use treatments. For one-time treatments, we assume that health benefits may last as much as a lifetime or the average duration of a disease. When the treatment lasts for more than a year, we annualize benefits and attribute one year of health improvements to the respective one year of treatment.

- Many health evaluation studies used in Constellation's metrics use randomized control trials and report direct impact of treatment on QALYs. Constellation always assesses the external validity of the experimental design and whether its conclusions are reasonably applicable to low-income individuals within the Twin Cities.

11.3 Procedures for the Use of Quality-Adjusted Life Years

Constellation uses QALYs as our preferred measure of economic value for health outcomes. QALYs are a measure of health output that can simultaneously capture gains from reduced morbidity (quality gains) and reduced mortality (quantity) gains, integrating these into a single measure (Drummond, et al., 2015). Constellation assigns a $50,000 monetary value per QALY to estimate benefits from improved health, which is the most commonly used value throughout the health literature (Neumann, Cohen & Weinstein, 2014). In the case of Constellation's metric system, we use this single value of QALY to compare all interventions.

All QALY and health-related benefits are discounted to present value. In many cases, QALY gains are already discounted when reported in the literature. When they are not already discounted, Constellation does so at a rate of 3% per year. When evidence suggests a lump QALY gain that in actuality will be realized in small amounts over multiple years, we assume a relevant time horizon (either suggested by the evidence as the duration of benefits or the duration from the participant's age at time of treatment to the life expectancy age of the relevant population) and assume that the total economic benefit estimated materializes in equal annual amounts during the time horizon. For example, if an intervention adds $5,000 of health benefits (a 0.1 QALY increase) for participants at an average age of 30 years with a life expectancy of 75 years, Constellation assumes that those benefits would accrue in equal annual amounts of $111 each year for 45 years.

Converting DALYs to QALYs

Health studies often report outcomes in terms of disability-adjusted life years (DALYs). DALYs estimate the burden of a disease or health condition using a "disability weight," or a value assigned to various conditions indicating the detriment to quality of life caused by the condition. DALYs can be converted to QALYs by dividing the DALY gain by a conversion factor, which is determined based on both the age of participants and the expected duration of the condition (Sassi, 2006). Conversion factors can be found in table 1 of Sassi (2006).

Constellation uses this process to convert from DALYs to QALYs. In the case where DALYs are not given but the change in disability weight is known, Constellation calculates DALYs by multiplying the difference in disability weight by the expected duration in years that the difference is expected to persist, and discounting to present value. DALYs calculated in this way can then be converted to QALYs.

QALY References

Sassi, F. (2006). Calculating QALYs, comparing QALY and DALY calculations. *Health Policy and Planning*, *21*(5), 402–408. https://doi.org/10.1093/heapol/czl018

Neumann, P. J., Cohen, J. T., & Weinstein, M. C. (2014). Updating cost-effectiveness: The curious resilience of the $50,000-per-QALY threshold. The New England Journal of Medicine, 371(9). doi.org/10.1056/NEJMp1405158

Drummond, M.F.; Sculpher, M.J.; Claxton, K.; Stoddart, G.L.; Torrance, G.W. (2015). Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Oxford : Oxford: Oxford University Press, 2015.

12. Employment Programs, Methods, and Assumptions

12.1. General Approach to Assess Employment and Economic Development Interventions

Constellation determines the method for assessing employment programs on a case-by-case basis. The following general assumptions are the basic starting point for any analysis.

12.2. General Assumptions

- The number of years of benefits (increased annual earnings) after an employment training period is determined by the evidence from program data or literature and usually covers between 6 months and 5 years (Council of Economic Advisers, 2016). Our preferred approach is to analyze the labor market prospects for each career or type of certification the program provides and establish a duration of benefits closer to the specific realities of the respective field. However, in some cases this tailored analysis is not possible, so instead we draw on categories described in Card, Kluve, & Weber (2017). These categories allow us to classify short-term impacts of less than a year post-program, medium-term impacts of 1-2 years post-program, and long-term impacts of more than 2 years post-program.

- We use pre-training earnings to approximate the amount of earnings in the absence of the program or the counterfactual earnings. Whenever that figure includes individuals with zero earnings, it concurrently measures the chance of having a job.

- To determine average annual post-program earnings, we must consider potential issues with data reported by programs, including bias or error from self-reported earnings and spotty or missing data.

- Regarding an alternative measure of the chance of having a job post-program in the absence of the program, Heinrich, Mueser, Troske, Jeon, & Kahvecioglu (2013) estimate that the probability of finding a job for participants in public job training programs included in the Work Investment Act (WIA) is between 50% and 60%.

Employment References

Card, D., Kluve, J., & Weber, A. (2017). *What works? A meta-analysis of recent active labor market program evaluations*. National Bureau of Economic Research. Working Paper 21431. Retrieved from: http://www.nber.org/papers/w21431

Council of Economic Advisers. (2016). *Active labor market policies: Theory and evidence for what works*. [Issue Brief]. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/page/files/20161220_active_labor_market_policies_issue_brief_cea.pdf

Heinrich, C. J., Mueser, P. R., Troske, K. R., Jeon, K.S., & Kahvecioglu, D. C. (2013). *Do public employment and training programs work?* IZA Journal of Labor Economics, 2(1), 6. doi.org/10.1186/2193-8997-2-6

Appendix: Additional Resources for Research

CDC	CDC - BRFSS Prevalence Data & Data Analysis Tools
Social Value UK	Resource Library - Social Value UK
Cochrane Training	Cochrane Handbook for Systematic Reviews of Interventions
MMB Results First	Minnesota Results First / Minnesota Management and Budget (MMB)
HHS Surgeon General	Reports and Publications | HHS.gov
IES - National Center for Education Eval and Regional Assistance (What Works Clearinghouse)	WWC | Find What Works!
US Department of Labor (CLEAR)	US Department of Labor (CLEAR)
Policy Insights Data	A Unified Welfare Analysis of Government Policies
US HHS, Administration for Children and Families (Prevention Services Clearinghouse)	Prevention Services Clearinghouse
Council of State Governments (What Works in Reentry Clearinghouse)	What Works in Reentry Clearinghouse
California Department of Social Services (California Evidence-Based Clearinghouse for Child Welfare)	The California Evidence-Based Clearinghouse for Child Welfare
NBER Health	The Bulletin on Health Archives | NBER
Center for Benefit-Cost Studies of Education	Center for Benefit-Cost Studies of Education
Tufts Cost Effectiveness Analysis Registry (health)	Cost-Effectiveness Analysis (CEA) Registry