



GLOBAL
INNOVATION
FUND

Practical Impact: GIF's approach to impact measurement



V1.01 June 2019

© Global Innovation Fund 2019



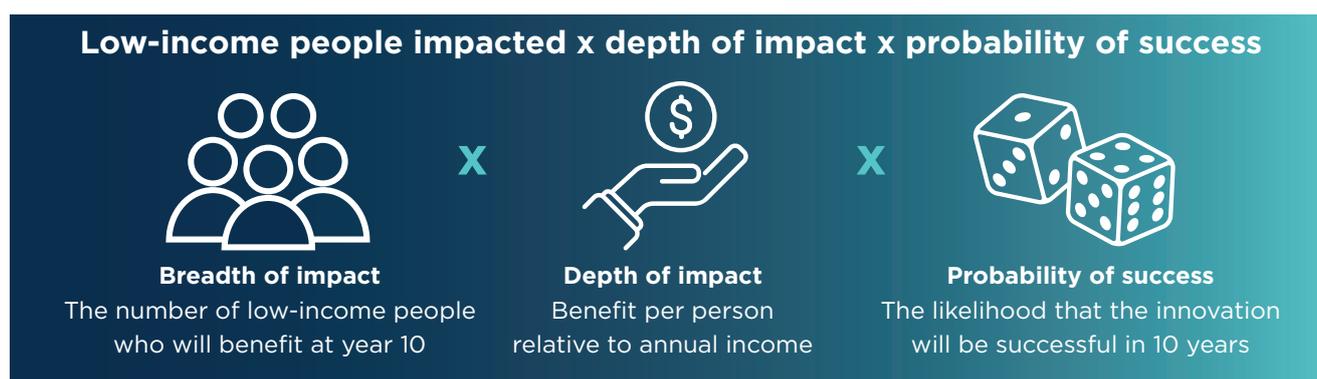
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Summary: Practical Impact in a nutshell

GIF seeks to maximize the social impact of its investments in innovations by using disciplined methods to:

- Forecast the impact of prospective investments and use this information to guide investment decisions.
- Track project performance and impact during implementation, using real time information to adapt and adjust as necessary.
- Evaluate investments after their completion to better understand how investments fared (and why), using this evidence to guide future GIF decisions; and inform decisions made by other development partners.

Practical Impact provides a structured way of forecasting the long-term impacts of early-stage innovations. During project selection and diligence, GIF computes Practical Impact as:



Breadth is the number of low-income people projected to benefit 10 years after the initial investment. The ten-year horizon is chosen as an index of the progress towards long-run maximum scale of the innovation.

Depth of impact is a universal measure that covers economic and noneconomic benefits. It captures relative changes in consumption or standard of living.

Probability of success is assessed by considering the risks along the path to scale.

Think of one Practical Impact unit as meaning: one person got a one-time benefit equal to 100% of their annual income (or consumption). We call this unit a person-year of income-equivalent (PYI). If 20 people each received a benefit equivalent to 5% of their annual consumption, that would also be reckoned as one PYI.

Practical Impact can be rapidly estimated to an order of magnitude during investment screening. During diligence, the estimates are refined. GIF's strong emphasis on evidence generation means that many investments will incorporate rigorous measurements of depth and breadth. This allows Practical Impact to be updated over time as risks are resolved and benefits are more accurately measured.

Practical Impact forecasts are expressed as a range, acknowledging the inherent uncertainty in forecasts. Because there is a common unit of measurement, these forecasts can be summed across the portfolio. At the portfolio level, these uncertainties average out to some extent. Risks to overall portfolio impact can then be assessed using simulation methods.

Table of contents

Table of contents	ii	3. Practical Impact during the investment life cycle	10
Glossary	iii	Selection and screening	10
1. Introduction	1	Diligence	10
Purpose of this note	1	During implementation	10
Background: about GIF	1	Portfolio monitoring	10
How impact assessment helps GIF maximize its social benefits	1	Evaluation and learning	11
Challenges in measuring impact	1	Mitigating optimism bias in Practical Impact estimates	11
How Practical Impact and social returns address challenges in measuring impact	2	4. Social returns	12
2. Basics of Practical Impact	3	Defining social returns	12
Practicality: order of magnitude is often good enough	3	Calculating social returns	12
Component 1: breadth: low-income people impacted	4	Practical Impact and social return: complementary impact tools	13
Component 2: depth of impact per person	4	Summary: social return versus Practical Impact	14
Measuring benefits: a relative scale	4	5. Conclusion	14
Netting out the cost of providing the benefit	6	6. Technical appendix	15
Component 3: probability of impact	6	Choice of discount rate	15
Understanding Practical Impact units	7	Normative approach	15
Practical Impact forecasts are snapshots	8	Opportunity cost approach	15
Practical Impact counts people impacted, not 'reached'	9	Conclusion	16
Uses and limitations of Practical Impact	9	Valuing mortality reduction	16
Cost-effectiveness of investment and attribution of impact	9	Valuing morbidity reduction	17
		Valuing education	18
		Valuing externalities of greenhouse gas increases or reductions	19
		Valuing domestic resource mobilization	20

Glossary

CPIA	Country Policy and Institutional Assessment
DALY	Disability-adjusted life year
ERR	Economic rate of return
GHG	Greenhouse gas
GIF	Global Innovation Fund
IDA	International Development Association
IEG	Independent Evaluation Group of the World Bank Group
IFC	International Finance Corporation
IRR	Internal rate of return
LSMS	Living Standards Measurement Survey
MCC	Millennium Challenge Corporation
NPV	Net present value
PYI	Person-year of income equivalent
PPP	Purchasing power parity
QALY	Quality-adjusted life year
RCT	Randomized controlled trial
SROI	Social return on investment
WTP	Willingness to pay

1. Introduction

Purpose of this note

This note provides a popular overview of why and how GIF forecasts and measures impact. GIF's unique Practical Impact methodology is central to the way GIF pursues its mission. We hope this introduction to Practical Impact will help donors, investees, and the public to better understand how we work. We hope also that other development financiers and impact investors may be interested in learning how GIF has addressed the challenge of investing for impact in innovations. This is a living document and will be updated as the methodology is further refined and extended.

Background: about GIF

GIF is a non-profit investment vehicle which invests in innovations to improve the lives of low-income people. GIF takes a venture capital approach – not to maximize its own profits, but to maximize the social benefits it creates. GIF does this by investing in early-stage innovations that are risky but which, if successful, offer prospects of large benefits to millions of people.

GIF mitigates risk by tying funding to evidence of the potential for impact. The stronger the evidence, the larger the commitment. Untested ideas get modest funding to demonstrate proof of concept. Where there is already a proof of concept, funding notches up. At this level, GIF supports its investees to build evidence on impact, cost-effectiveness, and market acceptance. With deeper understanding of impact and favorable conditions for scale-up, GIF is willing to invest more still.

How impact assessment helps GIF maximize its social benefits

GIF's funds are tiny, compared to development needs. With these limited resources, GIF wants to catalyze the highest possible social benefits. To that end, GIF uses impact analysis:

- during project selection and diligence, to identify projects with the greatest potential
- during project execution, to monitor and improve performance
- after project completion, to provide evidence on the cost-effectiveness or social returns of innovations for the benefit of emulators or follow-on investors
- at the corporate level, to assess GIF's performance in meeting its goal of creating social benefit

Challenges in measuring impact

GIF's Practical Impact methodology was devised to address three challenges.

Finding a common measure for impact across sectors

GIF supports investments across disparate sectors, including agriculture, health and education. In allocating its limited resources, GIF needs to make tough choices. For instance, putting time and money into improving literacy, versus saving lives, versus boosting farmer incomes. How can these options be compared? How can they be summed up for a portfolio view of impact?

Standard measures have shortcomings. 'People reached' does not distinguish between those who heard a radio message and those whose lives were saved. 'Lives improved' lumps minor improvements together with profound ones. 'Dollars spent' does not capture actual impact; and

might even reward inefficiency by making expensive programs look better than thrifty ones. Social returns approaches (project economic analysis, cost-benefit analysis, etc.), which monetize all costs and benefits, are theoretically attractive but highly demanding of time and information to compute.

Taking a long-run view of impact while meeting a demand for results accounting

GIF's investments are typically small, but they aim to catalyze growth to scale over a decade or more. This poses an accountability dilemma. Long-term impact is what counts in the end. But near-term impact is what can be easily measured today. A results framework that focused only on today's demonstration project would entirely miss the promise of vastly greater future impact. How can a results framework credibly report on progress towards that long-term impact?

Taking a portfolio-level view of goal-setting

Investing in innovations requires a portfolio approach. Although innovations are risky, GIF invests in them because it considers that the rewards justify the risks. A few flourishing innovations can outweigh those that fizzle. By taking a portfolio view, GIF remains fixed on maximizing total impact. We believe that this avoids the pitfalls that come from frameworks that tally the number of successes and failures, which can lead to excessive risk aversion.

How Practical Impact and social returns address challenges in measuring impact

GIF meets these challenges with two approaches to assessing impact. Each summarizes impact into a single measure that can be compared and summed across disparate project types:

Practical Impact is GIF's primary measure and is applied to all investments. Practical Impact provides an index of long-term benefits that might arise from a current investment. It can be updated as risks are resolved and impacts better understood. Practical Impact should not be thought of as a full social returns measure, but rather as a superior alternative to the 'people reached' metric.

Social returns assessments are applied selectively to larger and more mature investments. Social returns are measured in monetary equivalent terms, as the net present value of total benefits less total costs. This requires detailed information on the time paths of cost and benefits.

2. Basics of Practical Impact

Practical Impact serves three purposes related to GIF’s aim to maximize the social impact of its investments:

- A framework for initial screening: is funding justified, given the potential scale; impact; and likelihood of success?
- An updatable metric for projecting impact of projects under implementation.
- A framework for assessing the expected impact of GIF’s portfolio.

Practical Impact is designed to provide useful information while ensuring that the demands placed on staff and investees are reasonable and proportionate.

We estimate expected social benefits as:

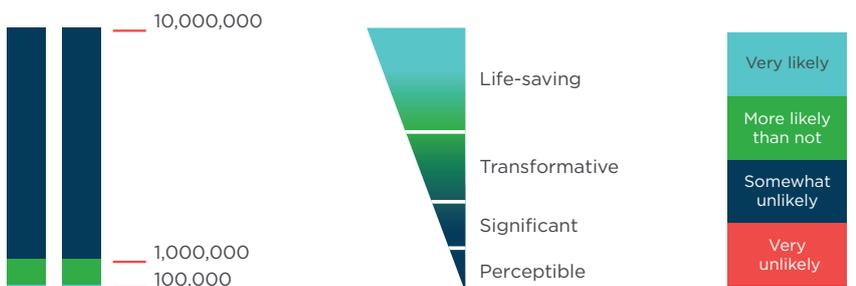


Figure 1: Practical Impact defined

Practicality: order of magnitude is often good enough

For quick appraisal, order of magnitude estimates are handy. Will the innovation benefit 100,000 people, one million, ten million or one hundred million? Will it make a 1% difference to their standard of living or a 10% difference? Is the chance of achieving

scale a long shot (10% chance)? Or a good bet (60% chance)? These will be judgment calls, but Practical Impact gives a structured framework on which to hang these initial judgments. As better information becomes available, these estimates can be refined.



How many people will benefit?

How big is the benefit per person?

What is the chance of success?

Figure 2: Order of magnitude: often good enough

Component 1: Breadth: low-income people impacted

Breadth of impact is assessed as follows:

- People impacted are those who are better off because of the innovation – this differs from ‘people reached’: a vaccination project may reach (vaccinate) 100,000 people, but people impacted are those whose lives were saved, or health improved, because of the vaccination.
- Practical Impact only tallies beneficiaries among GIF’s target population: those living on less than \$5 Purchasing Power Parity (PPP)/day.
- The focus is on the innovation, rather than the current investment. Breadth includes people impacted by follow-on investments and by spontaneous replication (this includes replication in other countries).
- Practical Impact is forecast, by default, to an anticipated level of scale-up achieved 10 years after the start of the current stage – it is a snapshot of impact in that 10th year, not a cumulative measure.

Component 2: Depth of impact per person

Practical Impact puts permanent and temporary income gains, health improvements and educational improvements on a common scale. This scale embodies value judgments. But so does any approach to quantifying impacts. For instance, measuring the benefits of life-saving interventions as ‘years of life saved’ versus ‘number of lives saved’ places a much higher value on saving a newborn’s life than that of an elderly person. Practical Impact makes the basis of comparison explicit and transparent. Recognizing the limitations of this approach, GIF uses Practical Impact as a starting point for discussion, not as a rigid basis for decisions about project selection.

Measuring benefits: a relative scale

For ease of application, Practical Impact is expressed relative to the beneficiary’s consumption level. This is convenient for the analyst, who might have a better intuition about relative benefits than dollar benefits. It might approximate the beneficiary’s internal yardstick for gauging the size of the benefit; and it deliberately focuses GIF’s attention on those with the lowest incomes, since a \$1/day increase will register as more important for the beneficiary living on \$2/day versus one living on \$5/day.

To build up the framework, we start with benefits that can be measured in monetary terms: changes in income or in the value of consumption. Practical Impact asks: is such a benefit just perceptible – a 1% to 3% permanent increase in consumption? Is it substantial – a 10% to 30% permanent increase in wellbeing? Or transformative – doubling the beneficiary’s consumption level for life?

Other benefits, such as health and education, can then be benchmarked using this ladder of depth scores (Figure 3 and Table 1). Depth scores for education are based on the relationship between schooling and wages (recognizing that education has other important and less quantifiable benefits). Health benefits are translated using methods from health economics (see Technical Appendix). Other benefits can be interpolated using the perceptible/substantial/transformational scale as a qualitative guide.

It is possible for a project to have multiple channels of impact, with a different breadth and depth for each channel. For instance, a clean water project may save lives, prevent disease and save time in fetching water. Practical Impact is summed across these channels.

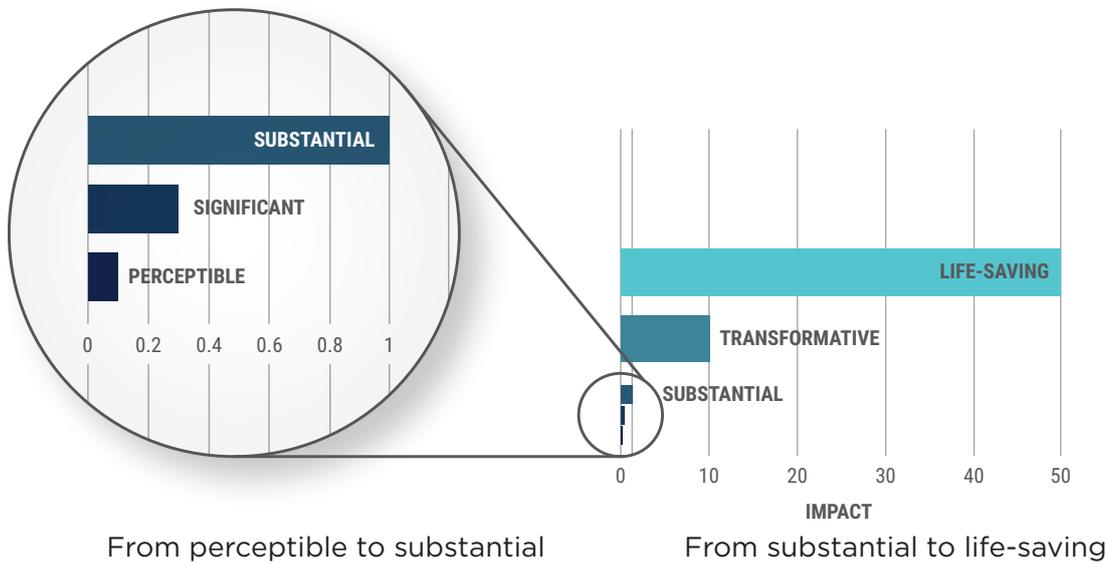


Figure 3: Depth of impact illustrated

	Depth factor	One-time benefits	Enduring benefits	Health	Education
Perceptible	0.1 to 0.5 Example: Boosting a yield for one season	10%-50% of annual consumption or equivalent willingness to pay (WTP)	1%-5% permanent impact on consumption or equivalent WTP		
Significant	1.0 Example: One year of education	100% of annual consumption	10% permanent impact on consumption or equivalent WTP		An additional year of primary education (0.5-1.75 – higher for girls and/or in Sub-Saharan Africa); and for higher quality of education
Substantial	5.0	500% of annual consumption	50% permanent impact on consumption or equivalent WTP	One disability-adjusted life year (DALY) averted	
Transformative	10.0		100% permanent impact on consumption or equivalent WTP	Preventing severe disability	
Life-saving	50.0 Example: A life saved via immunization			Avoiding a death	
<p>Note: The conversion between a one-time benefit and an ongoing benefit reflects GIF's discount rate of 10%. This means that a one-time benefit of \$100 is equivalent in value to a perpetual stream of \$10/year. For more information on this and on the health and education values, see the Technical Appendix.</p>					

Table 1: Quantifying depth of impact

Netting out the cost of providing the benefit

Innovations differ widely in cost-effectiveness. For instance, anti-poverty ‘graduation’ programs can boost income, but training and equipping the beneficiaries is relatively expensive. Conversely, behavioural nudges that increase people’s success at school, work or health practices may confer small but meaningful benefits at relatively small cost. Software platform innovations also have the potential to generate benefits at very low average cost per beneficiary.

Practical Impact nets out the average cost of providing the benefit, with costs computed to PYI equivalents. In practice, because GIF seeks innovations that are easily scalable and highly cost-effective, we hope that the difference between gross and net depth will usually be small for the innovations that make it into our portfolio.

Component 3: Probability of impact

This is the probability that the innovation achieves scale, reaching the anticipated year-10 breadth and depth of impact. Probability of success is the converse of risk, with probability estimates necessarily educated guesses, based on identifying and assessing critical risks. Research has shown that it is possible and useful to assign numerical probabilities to risks.

It is useful to break down the probability of success by stage of development. A stylized risk analysis looks like this:

Stage 1 (current): GIF’s current investment tests whether an innovation can achieve a critical step towards scale. For example, a grant might fund a rigorous test of an innovation’s cost-effectiveness. This could be a condition for adoption by a government agency. For private sector firms, it could be achievement of customer, revenue or cost benchmarks needed to attract a follow-on financing round. Achieving these critical steps is not a sure thing.

Stage 2 (scale): Success at stage 1 may be necessary, but not sufficient for the follow-on action. The government may decline to adopt the innovation; the funding round may not materialize. Even if there is follow-on support, the innovation may simply not work at scale.

Then:

Probability of success at scale = probability of success at stage 1 x probability of success at stage 2, assuming stage 1 success.

This captures GIF’s approach of scaling up an innovation through a series of tests: from proof of concept, to proof of cost-effectiveness, to proof of replicability. Different risk factors apply at each pass-fail juncture. As innovations are tested, risks are resolved one way or another and the probability of achieving scale can be updated.

Figure 4 illustrates an example, based on a GIF-supported project in Burkina Faso. The investment rigorously tests, through an experiment, whether saturation radio broadcast of a carefully-crafted message can boost uptake of modern contraceptives.

Demonstration stage risks:

- Does the innovation work?
- Is the evaluation adequately powered?
- Was the campaign executed as planned?

Scale stage risks:

- Potential support for program?
- Is there sustainable financing?
- Was the campaign executed as planned?

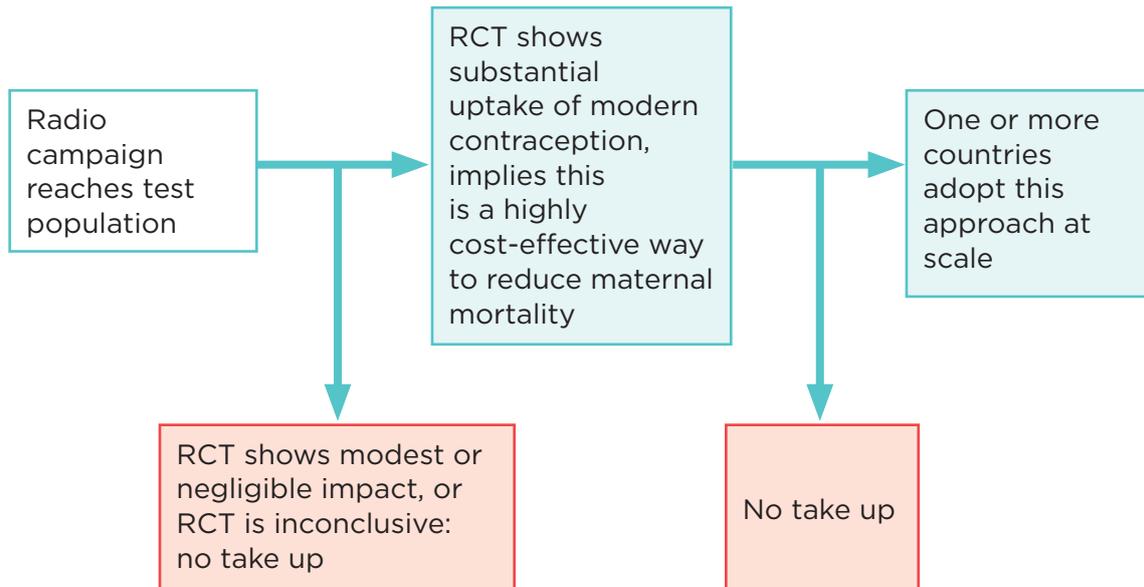


Figure 4: Risk tree for innovation scaling

Understanding Practical Impact units

Practical Impact is measured as people x depth of impact. Depth of impact is the benefit, measured in annual income (or consumption) per capita of the beneficiary. Think of one Practical Impact unit as meaning: one person got a one-time benefit equal to 100% of their annual income (or consumption). We call this unit a person year of income-equivalent (PYI).

So, one person getting a one-time gain of 15% of their income translates to 0.15 PYI. 100 people receiving this gain yields a total PYI of breadth x depth = 100 x 0.15 = 15.

Some interventions yield a time-limited benefit. Others, such as education or vaccination, yield an ongoing, indefinite flow

of benefits. To put these interventions on the same footing, we discount the value of future benefits. Because GIF uses a 10% discount rate (see Technical Appendix) there is a simple translation between flow benefits and one-time benefits. A flow benefit of x PYI per year is equivalent¹ to a one-time benefit of 10x PYI.

For example, suppose a person receives training in 2020 that boosts their annual income by 15% each year for the rest of their working life. Because we discount the future benefits using a 10% discount rate, this works out to a benefit of roughly 1.5 PYI. That benefit is 'booked' as occurring in 2020.

¹ This is an overestimate, but it is quite close if the flow goes on for 25 years or more – as is the case for education and vaccinations. If the discount rate is 0.1, a PYI next year is 'worth' $1 / (1 + 0.1)$ PYI today; a PYI two years from now is worth $1 / (1 + 0.1)^2$ and so on. As a result, the flow of x PYI per year is valued at $10x = x * (1 + 0.1)^{-1} + x * (1 + 0.1)^{-2} + \dots$

The following would all have a PYI measure of one million:

- permanently boosting the income of one million people by 10%
- providing a one-time benefit of 100% of annual income to one million people
- permanently boosting the income of 10 million people by 1%
- saving 20,000 lives

As noted, equating these outcomes embodies specific, but transparent, value

judgments. This is at once the advantage and disadvantage of any index number system. It is shared by systems that measure impact in disability-adjusted life years, which readers may find a helpful analogy (see Box 1), or in Net Present Value.

Because Practical Impact depth is expressed relative to an individual's consumption, Practical Impact can also be expressed as relative gains in wellbeing. For instance, a project expected to impact one million people and to generate 200,000 PYI could be described as providing one million people a year with a gain of 20% in welfare.

To assess the global burden of disease across countries and illnesses, the World Bank and World Health Organization devised a metric – the disability-adjusted life year (DALY).² Their problem was to compare and sum burdens across a diverse range of illnesses: from malaria, to tuberculosis, to cancer, to clinical depression. Their solution was to assign a severity weight to each condition, ranging from 0 (no disability) to 1 (death). Each condition also has an average duration in years. The DALY for a condition is the duration of disability, weighted by severity, plus the years of life lost to premature death. This also imposes a value judgment: two conditions with the same DALY are considered to be equally burdensome. Although people may disagree with the implicit value judgments, DALYs have proved to be a convenient and widely used tool and are often used for cost-effectiveness comparisons in health. By analogy, a PYI can be thought of as an impact-adjusted life year.

Box 1: DALYs as an analogy to Practical Impact units

Practical Impact forecasts are snapshots

Practical Impact forecasts are a snapshot of impacts made during just one year, 10 years after the investment starts. They don't include the sum of projected benefits from now until then.

Why 10 years and why a snapshot? Innovations, public and private, typically follow an S curve of adoption. Growth is slow at first, as the innovations are tweaked and refined. It can take more than 10 years for an innovation to saturate its potential market. But it is generally too difficult to predict beyond a ten-year horizon. We focus on the ten-year point as a feasible index of long-term impact. Using a snapshot saves the need to trace out (using further assumptions) the precise path of growth from now until year 10.

Practical Impact counts people impacted, not 'reached'

Because Practical Impact focuses on actual impact – how many people are better off because of the project – its reckoning of people impacted will be fewer than a tally based on people 'reached'. Consider, for example, a GIF project that supports improved systems for child immunization in Pakistan. A 'people reached' approach would claim 100 to 400 times the impact of the Practical Impact approach (Table 2). While Practical Impact's reckoning of beneficiaries is lower, it attaches a very high depth of impact to those affected.

² World Bank, *World Development Report 1993: Investing in Health*, <https://openknowledge.worldbank.org/handle/10986/5976>

(1) Children immunized by program = people reached
(2) Deprived children (living on less than \$5/day) immunized by program = 95% of (1)
(3) Deprived children immunized by the program <i>whose life is saved</i> = people benefited = 2.75% of (2)
(4) Deprived children immunized by the program whose life is saved and <i>who would not have been immunized in the absence of the program</i> = 10% to 35% of (3) = people impacted = 0.26% to 0.90% of people reached

Table 2: People truly impacted are fewer than those ‘reached’ (numerical values are illustrative)

Uses and limitations of Practical Impact

Practical Impact is flexible but particularly applicable to early stage innovations (see Table 3).

Innovation stage	Applicability of Practical Impact
Conceptual: new technology, but use case is not clear	Nil or limited
Pilot: rough notion of addressable market and potential channels of impact	Usually applicable, if channels of impact are known
Demonstration: some evidence or basis for estimating impact depth; addressable market; risks to scale	Highly applicable
Mature: organization with five-year projections of costs, revenues, activities; and evidence relating activities to social impact	Applicable, useful for comparison and aggregation across the portfolio; also may be possible and useful to calculate Net Present Value of social benefits

Table 3: Applicability of Practical Impact by innovation stage

Practical Impact is not easily applied to innovations that provide important but difficult to quantify benefits, such as those that promote human rights, culture or biodiversity conservation.

Cost-effectiveness of investment and attribution of impact

How can GIF compare the cost-effectiveness of alternative investments? GIF is typically not the only investor in a venture; and sometimes it may be useful to allocate the investment’s anticipated impact among co-investors. A natural way to do this is to divide the forecast PYI by the total amount of current investment. This gives a PYI/\$ value, which can be compared across investments. For instance, a large investment may have the prospect of 500,000 PYI but requires \$20 million in current

investment. So each dollar GIF invests catalyzes an eventual expected .025 PYI. An alternative investment might carry a forecast of 80,000 PYI but requires only \$400,000 in current investment. That alternative offers .20 PYI/\$, a substantially higher impact per dollar than afforded by the larger investment.

This is just a starting point for comparing relative effectiveness of investments. An investment decision should also take additionality or leverage into account, at least in a qualitative sense. For instance, if GIF’s participation was crucial to crowding in other funders, then GIF’s impact per dollar would be greater than the simple PYI/\$ ratio. However, there is no easy way to quantify additionality.³ Additionality claims have to be built on credible descriptions of how an investor’s financial terms, technical support, diligence or reputation made a difference to the investee’s fate.

³ For a thoughtful discussion see Carter Patrick, Nicolas Van de Sijpe, and Raphael Calel. 2018. *The Elusive Quest for Additionality*. Center for Global Development Working Paper. No. 495.

3. Practical Impact during the investment life cycle

Practical Impact is used throughout the investment cycle.

Selection and screening

Practical Impact is used in a quick, ‘back-of-the-envelope’ way for preliminary screening of proposals. Following the approach of Figure 2, the screener asks: what is our best guess about the order of magnitude, the breadth and the depth? This identifies the candidates that have at least a prospect of being truly impactful.

Diligence

As diligence proceeds, Practical Impact is applied with increasing attention to detail. This begins as a rough sketch of the potential channels of impact. For each channel: does the theory of change make sense? What are plausible scenarios of breadth and depth? What are the risk factors affecting breadth and depth? What are the likely costs? Who will bear the costs? Who will enjoy the benefits?

Identifying the potential channels of impact⁴ guides diligence. The goal is to identify key parameters and assumptions that impact financial return estimates. This guides further investigation. By the end of the impact and financial appraisals, some parameters will be sharpened, some assumptions upheld or discarded. This allows refinement of the Practical Impact estimate. If key impact parameters are not well known, the Practical Impact estimates will have a wide band of uncertainty. GIF may then focus resources on measuring them during investment execution.

During implementation

Practical Impact is updated as implementation reveals new information; and investees can report on Practical Impact breadth already achieved. Implementation resolves some risks and may surface new ones, allowing the probability of success to be updated. Patterns of uptake and diffusion allow revisions of expected breadth of impact; and built-in impact evaluations yield more reliable estimates of impact depth or cost-effectiveness. Pilot projects may pivot to unforeseen new avenues of impact. Taking all of this into account, Practical Impact estimates can and should be updated on a periodic basis to remain maximally relevant.

Portfolio monitoring

Portfolios mitigate the risk not only of financial loss, but also of the failure to have impact. The portfolio view focuses attention on the overall fund impact – not on the success or failure of the individual, risky investments. As outcomes are aggregated over investments, individual forecast errors are averaged out. PYIs, the common currency of Practical Impact, make this possible.

Using Practical Impact, an impact portfolio is assessed like a financial portfolio. Practical Impact values of investments are summed up to get aggregate expected impact. We can also assess the likelihood that we will do better or worse than expected.

Practical Impact is uncertain for two reasons. First, there is genuine risk in whether an innovation will succeed or fail, akin to a roll of dice. Second, our estimates are judgmental and subject to error. We roughly quantify both these sources of uncertainty for each investment. For instance, since we are

⁴ For instance, a water supply project might produce benefits through better health, improved agricultural productivity, and reduced time fetching water. Each of these channels might have a different breadth, depth, and cost.

uncertain about the potential breadth of impact, we assign minimum and maximum values, and allow for the possibility of anything in between. We also judge the probability that each investment reaches its impact goal. Then we metaphorically roll the dice ten thousand times in a simulation. On each dice roll, each investment randomly succeeds or fails, based on the assumed probability of success. And each key parameter varies randomly between its maximum and minimum. So for each of the ten thousand simulations we project the overall portfolio impact. On this basis we can say: there is an 80% chance that the portfolio impact will exceed 2 million PYI per year, and a 20% chance it will exceed 8 million.

Evaluation and learning

Learning is central to the GIF approach. When an investment exits GIF wants to assess whether, from society's viewpoint, it is worth replicating or further scaling up. For test-and-transition or scale-level investments, information generated during execution can be used to generate retrospective (achieved)

Practical Impact estimates and to update long-run Practical Impact forecasts.

Mitigating optimism bias in Practical Impact estimates

Investors support projects they are enthusiastic about and enthusiasm can spill over into overly optimistic views of an investment's prospects. This is a risk for both financial and impact appraisal. Practical Impact aims to mitigate this risk in several ways.

First, GIF's internal processes promote discussion and testing of impact, financial assumptions, and forecasts.

Second, GIF intends to update Practical Impact forecasts as risks are resolved and impact is measured. This will correct individual project forecasts. Over time, it will provide feedback on overall forecast bias.

Finally, GIF intends to set up a system of independent external review for a sample of Practical Impact estimates.

4. Social returns

Defining social returns

The terms ‘social returns’ and ‘social return on investment (SROI)’ refer to several different metrics. (Box 2). The common thread is to reckon an innovation’s net benefit to society as whole, expressed in monetary terms. All benefits to all people are added up; all costs

to all people are deducted; and future sums are discounted back to today’s dollars. The methodology for calculating social returns comes from long-established literature on cost-benefit analysis and project economic analysis.⁵

Three metrics for social returns

A standard metric for social returns is:

(1) the net present value (NPV) of (benefits – costs) = $\sum (benefits_t - costs_t) / (1 + \delta)^t$ where δ is the discount rate and t is time

Sometimes SROI is literally interpreted as:

(2) an economic rate of return (ERR):
the discount rate at which the discounted sum of benefits and costs is zero.

ERR sometimes cannot be calculated – for instance, in the happy circumstance when benefits immediately exceed costs. This will be a feature of some of the most preferred innovations!

A sometimes-used alternative is to calculate the ratio:

(3) NPV(benefits)/NPV(costs)

This metric is ambiguous, because the result can be sensitive to whether an impact is recorded as having a benefit or reducing a cost.

Box 2: Three metrics for social returns

Calculating social returns

Social returns analysis is frequently applied to public sector projects and those sponsored by development agencies. Ideally, the analysis should provide detailed distributional information. It can answer questions such as: How do benefits and costs fall upon low-income people and on women? Does the innovation tend to reduce inequities? Among those with higher incomes, are there powerful interest groups who benefit or suffer; and how does this affect the political and social

sustainability of the innovation? Do people outside the project’s boundary benefit or do they bear costs? For instance, an upland agriculture project may protect valley-dwellers from floods, or it may pollute their water supply.

To apply this approach to private-sector-led innovations, it is useful to break out the net benefits to the innovating firm from the net benefits to the rest of society. The latter includes the net benefits to the firm’s customers, either as consumers or as producers in their own right.

⁵ In the area of development investments, seminal works include Squire L and Van der Tak H G, 1975. *Economic analysis of projects*. World Bank Publications. And Jenkins G P and Harberger A C, 1992. *Manual: Cost Benefit Analysis of Investment Decisions*. Harvard Institute for International Development.

Consumer surplus – the additional benefit that consumers receive beyond what they pay – is an important feature of the kind of private sector innovations that GIF supports. If successful, these innovations will transform markets. They may introduce products or services that were previously unavailable, or they may drastically reduce the price of obtaining a service. Such transformations generate social benefits by expanding markets to new purchasers.

As an example of this type of analysis, [Björkegren \(2018\)](#)⁶ looks at the social benefits of the expansion of the mobile phone network in Rwanda over 2005-2009. During this time the price of a call fell by 76% and the network's geographical coverage of the country expanded from 60% to 95%: a transformative change in access. Björkegren reckons a net social benefit of \$474-\$530 million (2005 dollars), of which 51% accrues to consumers, 35% to the network operators and 14% to the government.⁷

Practical Impact and social return: complementary impact tools

Practical Impact and social returns play complementary roles at GIF. Practical Impact resembles a social returns or net present value forecast, but its goals are different and its computational burden lighter.

Practical Impact is a measure of long-term catalytic impact. Its stylized assumption is that today's investment is crucial to the survival and long-term growth of the innovation. The PYI/\$ ratio is a leverage measure that captures the degree to which today's investment catalyzes that future impact. It should not be interpreted as a return on investment, because it does not account for the stream of costs and benefits along the way to that impact. Unlike most social return analyses, Practical Impact places much greater weight on benefits to low-income people.

Practical Impact is applicable to the widest range of GIF investments, allowing them to be compared and aggregated on a common footing. Consistent with GIF's current focus, Practical Impact's sweet spot is for appraising the potential of innovations that are early in development.

Social return is a comprehensive accounting of the total social costs incurred and benefits received; and more demanding of data and effort than Practical Impact. At GIF, it will be applied selectively for the following uses:

A retrospective social return assessment can inform decisions about scaling or replicating innovations

– this is particularly true for ventures that will scale through the public sector. For instance, a social return assessment of a district-level program could help provinces to decide whether it is worthwhile to replicate the program in other districts. A deeper understanding of who benefits from a program; and who bears costs, can be important in designing a strategy for scale-up.

A retrospective assessment is useful for GIF's accountability and learning.

For innovations that have reached sufficient scale, social return is a refined metric of impact that can be compared to a range of other interventions. For sufficiently mature private sector investments, it allows GIF to break down the total social return into the direct financial returns to investors (including GIF) and the net benefits to the rest of society.

For sufficiently mature innovations, it may be possible and useful to project social returns prospectively. This may be particularly worthwhile where the innovation's main objective is to save money through operational efficiencies.

6 Björkegren D. 2018. 'The adoption of network goods: evidence from the spread of mobile phones in Rwanda'. *Review of Economic Studies*. <https://doi.org/10.1093/restud/rdy024>

7 The consumer benefits are net of the cost of service and of the cost of handsets. The operator benefits are gross of operating and investment costs.

Summary: social return versus Practical Impact

Table 4 compares social returns and Practical Impact as applied within GIF.

	Practical Impact	Social returns
Applied to	All projects	Larger, more mature investments
Applies greater weight to benefits to low-income people?	Yes, benefits receive higher weights, the lower the beneficiaries' income level; summary measure doesn't include benefit to those living on more than \$5 PPP per day	No
Costs considered	Rough adjustment for average cost per beneficiary of provision	All costs, quantified
Performed prospectively	Yes	Selectively, information permitting
Performed retrospectively	Yes	Selectively, information permitting
Time period	Prospectively: snapshot at year 10 post-GIF investment Retrospectively: cumulative over the period post-GIF investment	Cumulative over life of innovation

Table 4: Differences between Practical Impact and social return in GIF practice

5. Conclusion

Practical Impact is a disciplined way of making, and updating, forecasts of impact. At first glance, it may seem audacious to try to forecast impact 10 years in advance. But venture capitalists do something quite similar when they attempt to assign financial value to start-ups or enterprises that could be comparable in scale, maturity and ambition to the innovations that GIF funds. The venture capitalists are trying to assign a value to a

company several years hence. That value, in turn, depends on the company's hypothetical subsequent growth many years into the future. Value depends also on allowing for the possibility that the venture fails. Valuation requires experience, information and judgment. It pays off in profits for the venture capitalist. Practical Impact, it is hoped, when applied with experience and judgment, will offer payoffs in social benefits.

6. Technical Appendix

Choice of discount rate

The discount rate is used to equate present versus future costs and benefits. We discuss two approaches to determining a discount rate for GIF.

Normative approach

The social rate of time preference r is a normative discount rate derived from the Ramsey equation as follows:

$$r = \rho + g\varepsilon$$

- ρ is the pure rate of time preference, from a social perspective – this incorporates sheer impatience and the chance that society will cease to exist in the future due to nuclear, climatic or pandemic catastrophe
- g is the growth rate of the economy
- ε is the marginal elasticity of utility with respect to income.

Each of the elements of r has been subject to philosophical and empirical debate.

The UK government guide to cost benefit analysis, the *Green Book*,⁸ suggests a ρ value of 1.5%. Others have argued that (catastrophe aside, and holding income constant), it is unethical to prefer present over future generations, and so argue for a ρ closer to 0.

The second term ($g\varepsilon$) discounts the future on the assumption is that we (or our children) will be wealthier in the future and therefore get less enjoyment (utility) out of a marginal dollar than we do now. Here g is the per capita growth rate of the economy and ε is the elasticity of utility with respect to income.

The *Green Book* suggests, for the UK, $g = 2\%$, $\varepsilon = 1$, but recommends that these values be re-examined for other countries. We suggest $g = 3.53\%$ as an average value for the growth rate of the low and lower-middle income countries on which GIF concentrates (based on mean growth rate of this group of countries over 2006-2015; China is excluded).

The utility elasticity ε measures the gain in happiness with a marginal increase in income. [Layard, Mayraz & Nickell \(2008\)](#)⁹ analyze data on happiness and income from a number of household surveys in mostly high-income countries. They report a mean $\varepsilon = 1.24$. [Claxton \(2018\)](#)¹⁰ notes the lack of relevant data in low-income countries; and suggests sensitivity analysis using values of ε up to 2. This could be interpreted as extreme aversion to very low-income levels.

Together, these parameters suggest $r = 6\%$ to 8.5% . For comparison of investments within country, the Ramsey formula might be further adjusted by country-specific growth rates and elasticity estimates. Also, there is an important line of argument for applying declining discount rates for investments with very long-lasting impacts ([Arrow et al. 2014](#)).¹¹ The *Green Book* advocates applying a declining discount rate for benefits and costs accruing more than 30 years in the future.

Opportunity cost approach

The opportunity cost approach asks: what is the cost of capital used for the investment? Typically, this approach looks to capital markets to establish a rate. At GIF, we take a different approach. GIF is currently (2019) funded predominantly from governments that are concerned with economic development. The opportunity cost of funding GIF to

8 HM Treasury. *The Green Book: Central Government Guidance on Appraisal and Evaluation*. 2018. London: HM Treasury.

9 Layard R, Mayraz G and Nickell S. 2008. 'The Marginal Utility of Income.' CEP Discussion Paper no. 784. Centre for Economic Performance. Revised January 2008.

10 Claxton, Karl. 2018. Accounting for the Timing of Costs and Benefits in the Evaluation of Projects Relevant to LMICs. Guidelines for Benefit-Cost Analysis Project. Working Paper No. 8. <https://sites.sph.harvard.edu/bcaguidelines/>

11 Arrow K, Cropper M, Gollier C, Groom B, Heal G, Newell R, Nordhaus W, Pindyck R, Pizer W, Portney P, Sterner T, Tol R, and Weitzman M. (2014). Should Governments Use a Declining Discount Rate in Project Analysis? *Review of Environmental Economics and Policy*, 8(2), 145-163.

promote economic development is the return to traditional development assistance. We consider two estimates of that return.

First, based on a set of macroeconomic simulations, [Arndt, Jones & Tarp \(2015\)](#)¹² estimate mean internal rate of return to foreign aid to be 11.24% ± 7.12%.

Second, the Millennium Challenge Corporation (MCC) is notable for undertaking and transparently posting detailed cost benefit analyses of all its projects. MCC reports the weighted mean estimated ERR of 76 projects at closeout (when costs are known but benefits still unfolding) to be 15.1% ([Ospina & Block 2016](#)).¹³ However, ERRs will not be calculated for one third of projects. Of these, 22 were completed but lacked ERRs because of: insufficient data, no evidence of impact, or the project being small. The other 15 were cancelled for reasons including insufficient progress to achieve targets, military coups and under-performing contractors. It may be presumed that the ERRs on the 37 non-reporting projects may have been low and possibly even negative. If the ERRs were known, inclusion of these projects would be expected to reduce the overall mean return.

Conclusion

Relying on the opportunity cost approach, GIF adopts a 10% discount rate for assessing, comparing and aggregating project results (MCC also uses this discount rate), applied across projects and countries. This rate is slightly higher than the upper range for a normative approach. Where possible, GIF will publish data to allow application of alternative discount rates by external analysts who use a different standard.

Valuing mortality reduction

Mapping of life-saving outcomes onto Practical Impact depth values is motivated by studies of the value of a statistical life (VSL). The VSL is not intended to represent the intrinsic value of life. Rather, it summarizes actual and stated trade-offs people make in choosing between money and small changes in mortal risk. For instance, people demand higher wages to take on dangerous jobs or pay higher rents for safer dwellings. As an example, if people are willing to pay \$1,000 to reduce their risk of death by one percentage point, then the VSL is \$100,000. This means that, on average, if 100 people made this choice, \$100,000 would be spent and one life saved.

Choice of a VSL necessarily involves a profound value judgment. However, VSLs are applied when this judgment has to be made, one way or another. Proponents of VSL argue that its use promotes transparency and consistency of choices. VSLs are commonly used, for instance, in cost-benefit analyses of public health and safety initiatives.

There are relatively few VSL studies and estimates for middle-income countries and fewer still for low-income countries. The most comprehensive and recent survey is that of [Narain & Sall \(2016\)](#),¹⁴ feeding into [World Bank and IHME \(2016\)](#).¹⁵ [Robinson \(2017\)](#)¹⁶ uses these surveys to impute a VSL/per capita income ratio of 51 for a low-income country with a Gross National Income per capita of \$1,026 (2015 USD). The imputation makes use of an assumption that higher income countries are willing to pay more for a statistical life. Using an assumed income elasticity for this willingness to pay,

12 Channing Arndt, Sam Jones, and Finn Tarp. 'What Is the Aggregate Economic Rate of Return to Foreign Aid?' *World Bank Econ Rev*, first published online July 28, 2015 [doi:10.1093/wber/lhv033](https://doi.org/10.1093/wber/lhv033)

13 S. Ospina and M. Block. 2016. 2015 Report on Closeout ERRs. <https://www.mcc.gov/resources/doc/report-2015-closeout-errs>

14 Narain, Urvashi & Sall, Chris. 2016. *Methodology for Valuing the Health Impacts of Air Pollution: Discussion of Challenges and Proposed Solutions*. Washington, DC: World Bank.

15 World Bank and IHME (Institute for Health Metrics and Evaluation) (2016). *The Cost of Air Pollution: Strengthening the Economic Case for Action*. Washington, DC: World Bank.

16 Robinson, Lisa A. 2017. Estimating the Values of Mortality Risk Reductions in Low- and Middle-Income Countries. *Journal of Benefit-Cost Analysis*, 2017, 1-10. [doi:10.1017/bca.2017.14](https://doi.org/10.1017/bca.2017.14)

VSLs from higher income countries are used to derive those for low and middle-income countries. Robinson reports also an estimate of VSL/per capita income of 167 per capita incomes based on [Viscusi & Masterman \(2017\)](#) for a low-income country.¹⁷ However, this study draws data primarily from developed country VSLs.

GIF will use a depth of 50 in its impact calculations; and will not vary this weighting based on country or on the age, gender or employment status of the people whose lives are saved.

Valuing morbidity reduction

Compared to VSL, valuing morbidity reduction is even more technically difficult and less well studied in low and middle-income countries. As a pragmatic approach, GIF uses the method of [Hirth et al. \(2000\)](#)¹⁸ to derive the value in PYI of a disability-adjusted life year (DALY) that is consistent with the Practical Impact factor for mortality

reduction.¹⁹ A health condition's DALY is calculated as the duration of the condition weighted by a severity of the condition, on a scale from 0 = no impairment to 1= full impairment or death. Quality-adjusted life years (QALYs), used by Hirth et al, represent quality of life year on the converse scale, from 1 = ideal to 0 = death.

Generalizing the method of Hirth et al. we are equating two different ways of assigning a PYI value to averting a death. The first, explained above, sets a value of 50 PYI, based on an appeal to the value of a statistical life. The second views averting a death as a sequence of life years saved, from the time of the avoided premature death up to the person's life expectancy (at the time of the avoided death). For each of those life-years saved, the person saves one DALY, since the severity of death = 1. We want to set a value of a DALY so that the value attached to this sequence of saved life years is the same as the value assigned to a saved life. As a result, we solve for DALY value in:

$$50 = \text{DALY value} * (\text{mean discounted lost life years})$$

Since lives are lost at all ages, we take a population-weighted average:

$$\text{mean discounted lost life years at death} = \sum_{AGE=0}^{AGE=85} \left[\text{proportion of deaths at AGE} * \sum_{t=1}^{t=LE(AGE)} (1+r)^{-t} \right]$$

where LE(AGE) is the life expectancy at AGE. We use death rates and life expectancy for least developed countries for 2015 from the [UN World Prospects for Population 2017](#).

17 Viscusi, W. Kip & Masterman, Clayton. 2017. Income Elasticities and Global Values of a Statistical Life. *Journal of Benefit-Cost Analysis*, 8(2).
 18 Hirth, R.A., Chernew, M.E., Miller, E., Fendrick, A.M. and Weissert, W.G. 2000. Willingness to pay for a quality-adjusted life year: in search of a standard. *Medical Decision Making*, 20(3), pp.332-342.
 19 These were not developed to be measures of welfare, as pointed out by Robinson and Hammitt. (Valuing Nonfatal Health Risk Reductions in Global Benefit-Cost Analysis. Guidelines for Benefit-Cost Analysis, Working Paper no. 2, October 2017 review draft). But there are no available comprehensive, rigorous measures derived in developing countries. For the rough purposes of Practical Impact, QALYs or DALYs capture the idea that good health is more valued than poor health, and will suffice.

The result is a DALY value of 5.2 PYI, which we round to 5. This estimate is reasonably robust, because with a discount rate of 10% the discounted sum of life years is between 10 and 11 for most of the population.

[Hammit & Haninger \(2017\)](#)²⁰ question the assumptions – built into the QALY or DALY measure – that people’s willingness to pay for morbidity reduction is directly proportional to both quality of life and duration of illness. Based on a stated preference survey conducted in the US, they found that avoiding a one-year illness at quality of life = 0.2 was only worth about twice as much as avoiding spending the same duration at the much health quality level of 0.8 (on a 0 to 1 scale). Avoiding a five-year illness was worth just 20% more than avoiding a one-year illness at the same quality of life. While it is worth exploring the proposition that the value of avoiding an illness may not be strictly proportional to severity and duration, these results are too counterintuitive to serve as a value function; not based on developing country data; and in any case are not easily implemented.

Therefore GIF adopts a Practical Impact valuation of a DALY as 5 per capita incomes. Operationally, the Practical Impact value for an averting condition with an impairment weight of SEVERITY and duration D (in years) is computed at the time of onset as:

$$\pi = \sum_{t=1}^D SEVERITY * (1+r)^{-t}$$

where the discount rate r is taken to be .10. For durations of less than a year, use D*SEVERITY.

A source for SEVERITY values is the [Global CEA DALY Calculator](#).

Valuing education

Education has many benefits, but the most prominent is its impact on earnings. This provides a natural linkage to the income-anchored Practical Impact scale. Where there are important indirect effects of education, such as on reduced fertility or child mortality, these can be reckoned in addition. This approach does not try to quantify the potentially important non-financial benefits of education.

Economics has a long tradition of quantifying the impact of education on earnings via the Mincerian equation:

$$\ln(\text{earnings}) = a + b(\text{years of schooling})$$

The estimated coefficient b is interpreted as the proportional increase in earnings associated with an additional year of education. It is usually expressed as a percentage and called the ‘return to education’, analogously to a financial return. Although it should theoretically not be interpreted causally – there are many potential sources of bias – in practice it is often used as an indication of the marginal impact of providing education.

Two recent papers have performed this analysis on large numbers of household data sets from the developing world. [Peet, Fink and Fawzi \(2015\)](#)²¹ used 61 consistently-formulated, high quality surveys²² from 25 countries (mostly in Sub-Saharan Africa and Latin America) over the period 1985 to 2010. They limit their analysis to workers with wage or salary income. They find that in the lowest-income countries (<\$500/year), the return to a year of primary education is 5.2%. For middle-income countries it is 6.2%. They find also that returns have declined over the past decade.

20 Hammit, J.K. and Haninger, K. 2017. Valuing nonfatal health risk as a function of illness severity and duration: Benefit transfer using QALYs. *Journal of Environmental Economics and Management*, 82, pp.17-38. <http://dx.doi.org/10.1016/j.jeem.2016.10.002>

21 Peet, Evan D., Günther Fink, and Wafaie Fawzi. ‘Returns to education in developing countries: Evidence from the living standards and measurement study surveys.’ *Economics of Education Review* 49 (2015): 69-90.

22 From the Living Standards Measurement Study (LSMS).

[Montenegro & Patrinos \(2014\)](#)²³ use harmonized surveys covering 819 economy-years in 139 economies, over 1970-2013 (with 96% after 1990 and 75% after 2000). They also restrict the sample to wage earners; and control for years of labor market experience. Table 5 summarizes their findings on returns to an additional year of primary education. A noteworthy finding is one of higher returns to female education.

The two studies give disparate results (perhaps due to different country samples) and both are subject to bias through exclusion of self-employed workers, among other reasons. But they motivate a Practical Impact depth of 1 ± 0.5 to a project which boosts education by a year; and of three to nine for a project which converts illiterates to literates (the equivalent of going from nought to six years of primary school). The choice of Practical Impact depth depends on the country, on gender focus and on the analyst's assessment of the potential impact of the educational intervention on educational quality.

Region	Female	Male
All (including high income)	13.2	10.1
South Asia	4.8	4.7
Sub-Saharan Africa	17.5	12.5

Table 5 Returns to a year of primary schooling, % (from Montenegro and Patrinos 2014)

Valuing externalities of greenhouse gas increases or reductions

Many projects will result in increases or decreases in greenhouse gas (GHG) emissions. Low-income people are highly vulnerable to climate change, with impacts already evident and due to intensify over coming decades. There is literature that attempts to quantify those costs by simulating different paths of emissions, development and climate change over the next century. These costs could then be translated into a carbon value (per ton of CO₂ emitted or avoided). However, these are complex exercises and may not well represent the risk of catastrophic impacts for which those with the lowest incomes would be least resilient.

GIF follows instead the approach and conclusions of the [High Level Commission on Carbon Prices](#).²⁴ The Commission sought to nominate carbon prices, which, if incorporated in the global economy today (e.g. via emissions taxes), would support achievement of the 2015 Paris Agreement to hold “the increase in the global average temperature to well below 2°C above pre-industrial levels and to pursue efforts to limit the temperature increase to 1.5°C”. The Commission concluded that to reach this goal, carbon prices should be at least US\$40-80/ton CO₂ and US\$50-100/ton by 2020. If we set the average value of a PYI as PPP \$1,000, then one ton of CO₂ mitigation represents a benefit of .04 to .08 PYI; one ton of emissions represents a decrement of benefits.

23 Montenegro, Claudio E. and Patrinos, Harry A., *Comparable Estimates of Returns to Schooling Around the World* (September 1, 2014). World Bank Policy Research Working Paper No. 7020.

24 <https://www.carbonpricingleadership.org/report-of-the-highlevel-commission-on-carbon-prices/>

Valuing domestic resource mobilization

Some projects aim to increase domestic resource mobilization or spur cost efficiencies in government, public utility or expenditure. Some of these new resources are assumed to be devoted to low-income people. These may be very diffuse benefits spread over a large population.

A practical approach is to take the amount of resources mobilized and estimate the proportion that will be devoted to low-income populations. For national budgets, for instance, the proportion devoted to primary health and primary education could be taken as a lower bound. Then the estimate of mobilized resources for the benefit of low-income people is divided by the nominal equivalent of one PYI for a representative low-income (under \$5 PPP) person. This gives the number of PYI generated, irrespective of the number of beneficiaries.