

Swarthmore College

Works

Economics Faculty Works

Economics

2018

Foreword: Why Measurement Of Costs And Benefits Matters For The SDG Campaign

S. Dercon

Stephen A. O'Connell

Swarthmore College, soconne1@swarthmore.edu

Follow this and additional works at: <https://works.swarthmore.edu/fac-economics>



Part of the [Economics Commons](#)

[Let us know how access to these works benefits you](#)

Recommended Citation

S. Dercon and Stephen A. O'Connell. (2018). "Foreword: Why Measurement Of Costs And Benefits Matters For The SDG Campaign". *Prioritizing Development: A Cost Benefit Analysis Of The United Nations' Sustainable Development Goals*. xxiv-xxx. DOI: 10.1017/9781108233767.001
<https://works.swarthmore.edu/fac-economics/464>

This work is brought to you for free by Swarthmore College Libraries' Works. It has been accepted for inclusion in Economics Faculty Works by an authorized administrator of Works. For more information, please contact myworks@swarthmore.edu.

Foreword

Why Measurement of Costs and Benefits Matters for the SDG Campaign

The Sustainable Development Goals (SDGs) provide an extraordinary vision of what global development should look like between now and 2030. Starting with the concept of sustainability, the SDGs go far beyond the Millennium Development Goals (MDGs) to incorporate a set of environmental and social-justice priorities that require national action at all levels of income. As agreed by 193 signatory nations at the September 2015 United Nations General Assembly, the 2030 Agenda for Sustainable Development (<https://sustainabledevelopment.un.org/post2015/transformingourworld>) is meant to be *universal*, *indivisible*, and *interlinked*. In conventional development arenas like extreme poverty and hunger the SDGs also inspire, doubling down on the MDGs by defining success in absolute rather than relative terms. Global partners target an end to poverty in all its forms, for example, rather than a 50 percent reduction in extreme-poverty head-count ratios.

The UN's drive for universal norms and targets involved widespread public debate and painstaking negotiations and compromises between national governments. The process was simultaneously more transparent and much more difficult and convoluted than when the MDGs emerged from behind closed doors a decade and a half ago. Some widening in the scope of commitments was inevitable and also desirable, to accommodate sustainability goals and build a truly global coalition. But there was also widespread awareness as negotiations proceeded that fewer goals might allow for greater success. By the latter standard, the 2030 Agenda is daunting. With 17 global goals and 169 highly ambitious targets, the Agenda seems in danger of departing not just in scope but also in coherence from the elegant eight goals and 17 targets of the MDGs.

In practice, therefore, a great deal remains on the table in terms of shaping global action. This is true not just in the conventional sense of identifying cost-effective approaches to individual targets but also in the deeper sense of operationalizing – and unavoidably, prioritizing – targets at the national and global levels. This book makes a vital contribution to what should be a collective effort to prioritize.

Cost-benefit analysis (CBA) is a well-established method for prioritizing spending in a world of limited budgets, not least in some of the poorest settings of the world. When done carefully, CBA and its cousin, cost-effectiveness analysis (which evaluates alternative approaches to achieving a given result), provide a transparent and evidence-based approach to identifying cost-effective uses of public money. Working together with ex-post evaluation and careful monitoring during program delivery, CBA can increase both the quality and the quantity of public spending, by shifting funds toward high-value projects and convincing funders (ultimately, taxpayers) that they are getting value for their money.

The Copenhagen Consensus should be applauded for its campaign to bring rigorous CBA evidence to bear in public debates on the scope of the SDGs. The papers collected here informed a comprehensive scorecard that covered the majority of the proposed targets and was available during the final year of negotiations. The analysis suggested what was at stake: assuming best-practice interventions, a failure to prioritize across goals could reduce a comprehensive measure of total benefits by 75 percent or more per dollar of costs. Losses of similar magnitude could accompany the pursuit of overambitious target levels or suboptimal interventions.

To date, this analysis has had less traction than the Copenhagen Consensus hoped, a result familiar

to any practitioner of cost-benefit analysis within governments and development agencies. The hope that the analysis would guide a winnowing of the goals did not materialize. But these studies remain crucial, as inputs into the debates that will now be required to operationalize the SDGs.

We focus on two questions in this foreword. First, why prioritize? We will discuss what the MDGs accomplished and how these lessons should inform the SDG process looking forward. Second, how should development actors – governments, development agencies, and nongovernmental and civil-society organizations – use the cost-benefit evidence collected here?

We should be candid at the outset on two matters. First, neither of us is convinced that top-down goal setting within international organizations represents the best route to development success. Development efforts require local and national political buy-in to be successful. Achieving growth and development in a society is complicated, messy, and context-specific and is about more than allocating resources. Outside of narrow corridors within which best practices are known and the links between inputs and outcomes are tight (as with some public health and humanitarian interventions), the allure of “buying” development – reducing development to spending a particular sum of money – is an illusion. Second, cost-benefit analysis has its own methodological limitations. Almost by definition, the clarity of a benefit-to-cost ratio is greater than what the data and modeling apparatus can support. Used uncritically, the method can support overconfident rankings between outcomes that are not easily compared and invite generalization across contexts that differ in unmeasured ways. Despite these observations, however, we strongly believe that cost-benefit exercises as conducted in this book should get more attention and should be used in debates on the allocation of resources across the world.

From MDGs to SDGs

The SDG process was spurred on by global successes during the MDG period, including a spectacular outcome for the extreme-poverty

headcount ratio that was confidently predictable well before the SDG consultations began in earnest. The argument for doubling down, however, rested on a claim that observed outcomes were the *result* of the MDGs, implying that they would not have occurred without the goals, targets, and institution building of the campaign. That claim remains controversial, for the simple reason that the counterfactual – the outcome that would have emerged without the MDGs – is not observable. Still, a few facts stand out that may well be attributable to some extent to the presence of the MDGs.

The global aid envelope expanded dramatically in the period since the MDGs were agreed, from US\$80bn in 2000 to US\$147bn in 2015 and after a period of stagnation in aid volumes during the 1990s. The clarity of the narrative around the MDGs may have helped to revive political interest in aid, amid fairly widespread disillusion among rich countries in the 1990s. The chosen goals were modest in number, and they were sufficiently non-controversial to mitigate conflicts of interest between donors and recipients. Their collective adoption was consistent with ongoing efforts to enhance donor coordination and avoid costly duplication of activity. Numerical targets were a key innovation of the MDG campaign: they promised an increase in two-way accountability, underpinned by credible and transparent mechanisms to monitor progress.

The discourse of *what gets measured gets done* acquired impetus late in the MDG campaign, reflecting a growing perception that the adoption of numerical targets did succeed in increasing accountability throughout the development cooperation system. By this argument, sending countries acquired leverage for holding recipients to account in the use of their funding, while recipient countries and other stakeholders were able to assess the alignment of donor portfolios – the countries and programs donors were willing to fund – with MDG priorities. Both sides plausibly faced new costs of renegeing on MDG-related commitments, as no stakeholder could publicly repudiate a target like cutting poverty in half.

The Department for International Development (DFID) and the United States Agency for International Development (USAID) both made

major efforts to bolster accountability during the 2000s, tied in some cases directly to MDG targets. USAID's Feed the Future program, for example, adopted the headcount ratio as a program target within its zones of influence, while DFID increasingly concentrated its spending in countries failing in income poverty reduction and the other MDGs. Late in the process both the World Bank (2013) and USAID (2014) appropriated extreme-poverty targets directly into their mission statements. The World Bank and UN system invested heavily from the outset in publicly available data and monitoring around the MDGs, an activity that undoubtedly spurred new research and may have facilitated watchdog innovations, including the Center for Global Development's aid-quality measures.

Formal attempts to construct a convincing counterfactual will continue. To date, the research has been limited to controlling for preexisting trajectories by looking for improvements in indicator trends among aid-receiving countries around the time the MDGs were adopted. Timing may of course be a weak proxy for the intensity of treatment, given that countries differed sharply in their exposure to MDG-related aid flows and that donor priorities had already moved decidedly in favor of poverty-reduction goals during the Highly Indebted Poor Countries Initiative of the late 1990s. These concerns notwithstanding, the research to date suggests a decidedly mixed picture: some indicators are consistent with a new departure around 2000, and others are not (see, for example, World Bank, Global Monitoring Report (GMR) 2016/2016). They also carry a sobering message looking forward because if preexisting trends represent a legitimate counterfactual, then the successes of the MDGs have made the remaining task considerably more difficult. The countries with the biggest indicator deficits in 2015 are, in many cases, those with the most adverse indicator trends over the past decade. This is in sharp contrast with China and India, which had the largest poverty deficits in 2000 but were already achieving spectacularly favorable (China) or at least modestly favorable (India) indicator trends before 2000. China, of course, received almost no development assistance

after 2000, and India received very little on a per capita basis.

Our own view is that whatever else the MDGs achieved, the campaign revitalized global development efforts by expanding aid flows and increasing accountability and coordination among donors. The troubling question is whether the sprawling scope of the SDGs puts these achievements at risk, especially against the headwinds of slower global growth. An agenda that is too broad to galvanize focused action may fail to sustain overall aid flows, misdirect such flows as are available, and risk returning the development community to a low-accountability mode of business as usual.

The SDG agreement shows clear if indirect awareness of this concern, pushing back vigorously with its characterization of the goals as universal, indivisible, and interlinked. From this perspective, the SDGs are less a set of competing goals than a comprehensive checklist for achieving the one great objective of ending global poverty on a sustainable basis. This interpretation is broadly consistent with the World Bank's interpretation of its own extreme-poverty mission (see World Bank, GMR 2015/2016, referenced earlier), and with USAID's *Vision for Ending Extreme Poverty*. These interpretations give targets for the extreme-poverty headcount ratio pride of place, but they define poverty as a multidimensional and contextualized phenomenon and lay out a theory of change that is broad enough to validate a very wide list of complementary targets.

But this returns us to prioritization. A central contribution of the MDG campaign was to elevate a plausibly universal concept of development itself – not as economic growth or progress, as crucial as those might be on instrumental grounds, but as elimination of human deprivation. The SDGs double down here as well, by incorporating sustainability and an insistence on *leaving nobody behind*. But characterizing a set of 169 targets as indivisible and interlinked comes close to repudiating any attempt to prioritize or assign responsibility. Accountability may lose its foothold if most forms of development spending can be validated in terms of their direct objectives while weak impacts can be explained away through appeals to inadequate efforts by other actors or failures

elsewhere in the system. And even where lines of responsibility are clear – as in the data-collection arena where the public-good aspect demands public provision – the magnitude of the task overwhelms available resources. In short, the leverage implied by *what gets measured gets done* strains credulity when stretched so far. We cannot credibly claim that *whatever* gets measured gets done.

Finally, we worry that a proliferation of targets may run afoul of some well-defined perils of scale. Numerical targets risk extending a gap-filling mentality beyond its appropriate domain. They perpetuate the impression that development outcomes can be purchased at a unit cost that is invariant across countries. They can enforce over-uniformity, favoring large-scale commitments that may stifle experimentation and fail to exploit individual-country or individual-donor opportunities. If these concerns vary in systematic ways across goals, the implication is that some goals lend themselves more readily to such targets than others.

How Should Development Stakeholders Use These CBAs?

The need for prioritization is clear in our view, and therefore the drive for sensible criteria to inform global debates. The chapters collected here provide benefit-to-cost ratios for a wide range of targets, assuming best-practice interventions. To interpret these ratios, consider an intervention that incurs an up-front cost of c to deliver a perpetual stream of benefits equal to b dollars each year (adjusted for inflation). Suppose that future costs and benefits are discounted at rate $r > 0$; the studies collected here compare 3 and 5 percent ($r = 0.03$ and $r = 0.05$). Then the ratio of discounted benefits to discounted costs – or *benefit-to-cost ratio* – for this intervention is given by $BCR = (1/r) * (b/c)$. This calculation illustrates the standard result that higher discount rates (embodying greater societal impatience) discourage interventions whose benefits are deferred relative to costs. At bottom, however, the intervention caricatured here provides discounted benefits of BCR dollars for every discounted dollar of cost. If a private firm could

recoup its costs by collecting a revenue stream equal to b each year, any intervention with a BCR exceeding one would be privately profitable. But in a social cost-benefit analysis the costs and benefits include environmental and third-party impacts that are not priced in markets, along with indirect impacts that may include synergies with other targets. Interventions that are socially profitable by a BCR criterion – even hugely so – typically require public intervention precisely because they are not privately profitable.

The difference between a target's BCR and 1, multiplied by the scale of the intervention, summarizes what happens to the total economic pie, including the valuation of goods and services that are not priced in markets, as a result of achieving the target (we emphasize scale effects later). The calculation is meant to be comprehensive, including all direct and indirect impacts. A BCR above 1 therefore means that the overall pie is bigger, and by a larger amount per dollar of cost the bigger is the BCR . In the absence of distributional weights (see later), an intervention with a BCR above 1 delivers enough dollar-equivalent gains per dollar of cost that nobody has to lose, at least in the hypothetical sense that a set of costless side payments would make it possible to fully compensate any losers while leaving at least one person better off.

Three key features shape these chapters and the resulting rankings sufficiently to warrant some general observations for nonspecialist readers. The first is the curse of diminishing returns. At the level of ambition embodied by the SDGs, a number of global targets (including those for global average temperature, primary and secondary enrollment, and maternal mortality) are subject to sharply increasing marginal costs. The cost of reducing projected global temperatures by 2 degrees over a given horizon, for example, is much more than twice the cost of reducing projected temperatures by 1 degree. In the presence of rising marginal costs, the best becomes the enemy of the good, and CBA has a natural tendency to produce moderation. $BCRs$ that are high at modest target levels start to fall as targets become more inspiring, and can go well below 1. The extreme-poverty headcount ratio falls to this

argument – getting to zero is too costly. This effect is even stronger if benefits are declining on the margin, but the curse discourages extreme targets even when goals are viewed as intrinsic rights that must ultimately be satisfied in full as rapidly as feasible.

In a world of diminishing returns, smaller interventions will tend (other things equal) to produce larger BCRs. The optimal set of interventions over any fixed overall budget and time horizon will therefore tend to involve the partial fulfillment of multiple targets. The argument for focusing on a few big efforts has to come from somewhere else – in short, either from a prioritization of rights that classifies selected targets as nonnegotiable, or from some form of increasing returns to individual targets. Our arguments about accountability fall into the latter category. They embody a form of increasing returns, where the cost of effective action includes a large fixed component that may involve data provision, coalition- and institution-building, or development of target-specific supply chains. These costs are implicit in the book, in the sense that all of the chapters take ambitious goals and large-scale efforts as a starting point. Other sources of increasing returns, including network effects (e.g., in stopping epidemics) and irreversibilities (e.g., in environmental preservation), play an important role in some of the relevant chapters. But the curse of decreasing returns inevitably pushes a number of authors to embrace more moderate target levels than the SDGs propose.

The second feature relates to the valuation of benefits. Within development agencies and governments, it is often sufficient to treat in-kind targets as given and focus on the search for cost-effective interventions. The chapters collected here perform a similar (and invaluable) task on the cost side – a task that is heroic enough on its own, given the unavoidable distortions of having to assume, first, that interventions at a given global scale encounter the same unit costs everywhere in the world and, second, that these costs can be reasonably estimated using one or two well-designed impact assessments from particular times and places.

But authors were also asked to place dollar-equivalent values on all benefits, so that users could compare global temperature targets with

completion of the Doha round and coral reef preservation with reductions in maternal mortality. Although expressing all benefits in dollar-equivalent values remains controversial, the appeal of this approach is obvious: if the analysis is even reasonably robust, it is hard to argue that projects with *phenomenal* BCRs (to use the Copenhagen Consensus's term for BCRs of 15 or above) should not receive priority relative to those with BCRs below 1. But the chapters vary widely in the comprehensiveness and robustness of their benefit estimates. Calculations of the social return to schooling, for example, are often famously modest in the sense of including only the social costs of schooling and none of the spillover benefits that a vast and admittedly contentious literature has emphasized over the years – spillovers that range from lower fertility to higher civic engagement and from improvements in institutional quality to women's empowerment and economy-wide innovation. Our own view is that these spillovers are of the essence. But Chapter 6 by Psacharopoulos is in this modest tradition – no spillover benefits, no synergies with other SDGs.

There may, in fact, be a general case for staying modest, given how contentious the assessment of these effects can be. And one does not need spillovers, for example, to favor a shift toward early-age interventions in education and health, given the increasing evidence of lifelong impacts on productivity and well-being. But the main point is caveat emptor: some chapters are braver (or more foolhardy) in this respect, and a more uniform treatment of benefits might substantially alter the rankings. The lesson is a general one when comparing CBAs across disparate sectors: users need to be attentive not only to how benefits are valued but also to what benefits are included.

The final feature relates to distributional objectives, which are central to the MDG and SDG campaigns but curiously absent in the cost-benefit calculation we described earlier. A thought experiment brings out the issue. Suppose for a moment that costless transfers were indeed possible and that the most cost-effective way to end extreme poverty was simply to guarantee each person on earth \$1.90 a day. This would be done through targeted transfers to make up any difference relative to each

person's market-related outcome. What BCR would this intervention generate? The answer is that unless the intervention altered the behavior of the household in some fundamentally favorable way – rather than just scaling up its consumption – the BCR could not exceed 1. The program benefit would be the discounted global consumption shortfall of the poor in the presence of the program – call this S – and the cost would be S as well. Any realistic accounting for administrative costs would in fact drive the BCR below 1.

Any outright efficiency gains from poverty reduction would help to push the BCR above 1. But some form of distributional weighting is arguably central to justifying any global poverty target. In the welfarist tradition within economics, this is done by making the social utility of an income-equivalent benefit depend on the household's income. A dollar of purchasing power is viewed as being worth more in the hands of a poor household than in the hands of a rich household. A rights-based approach has a similar feel: if \$1.90 is an absolute right, then only another right can be in tension with it, not a cost that may happen to exceed \$1.90.

Distributional concerns are handled in subtle ways in these chapters and readers should be prepared to query the individual chapters. In Chapter 24, on poverty, Gibson uses a modified version of S to measure costs. He assesses benefits, however, based on microeconomic evidence on the difference in lifetime earnings between individuals who grew up above and below the poverty line. This raises the BCR above 1, under the implicit assumption that some plausible combination of credit-market and information imperfections prevents the poor from borrowing to secure these efficiency gains themselves. But a simple distributional-weighting scheme could easily have raised the BCR of higher. Using log utility, for example, the value of transferring a dollar from a rich household to a poor household is not 1 but y_{Rich} / y_{Poor} . Logarithmic weights would therefore immediately translate an ambitious consumption-poverty target like 3 percent into phenomenal range because of its highly targeted beneficiary population (by implication, of course, the overall size of the pie is no longer the optimality criterion).

An implicit form of distributional weighting is embedded in some of these chapters, as when researchers apply an economy-wide value for disability-adjusted life years in evaluating health interventions that disproportionately favor poor communities. In these cases, as with distributional weights, the analyst places greater value on the well-being of the poor than their own willingness to pay would be able to reveal.

With these observations in mind, these chapters and the resulting rankings deserve a broad readership among development stakeholders and will raise the equality of public debates on priorities.

There are challenges and debates here for researchers as well. How far can a CBA platform take us in comparing health interventions with education interventions, let alone in accommodating improvements in accountability or sustainability? Can increasing returns and distributional impacts be handled more systematically? Is there external validity in the cost and benefit data, so that BCRs based on exemplary microeconomic evidence from individual countries can be appropriated for global calculations? Or do we actually have enough data to disaggregate in some cases – for example, to settle the costs of delivering a nutrition program in South Sudan, versus in Peru or India, all places with considerable stunting? How about synergies and general-equilibrium impacts; in some cases these are intrinsic to the calculation, as in the case of trade-policy reforms, while in others they are brought in selectively, as in the case of family-planning interventions that generate positive externalities through slower population growth. In still other cases they are excluded as too speculative. How important are these differences, and are there ways to formally incorporate successively more speculative elements of the analysis? Finally, how should the empirical methods employed to estimate treatment effects affect the interpretation of results? Should estimates based largely on randomized controlled trials, for example, be viewed as inherently conservative, while those that rely mainly on cross-country empirics or simulation modeling are viewed as decidedly less so?

Caveats are easy – too easy, because those who find these calculations uncomfortable will want to

dismiss them. We ourselves would not recommend spending the global development budget, or even the portion allocated by foreign aid agencies, simply based on the benefit-cost ratios in this book. But the contributions here are nonetheless invaluable. By providing a rigorous examination of the cost and benefit evidence, they are a crucial buttress to the morally urgent work ahead. They ask an unavoidable question: when resources to improve the lives of the poor are scarce, how can we get these resources to go further – much further? The question is difficult, but cost-benefit analysis provides a set of answers that are transparent and evidence based. Their transparency favors debate and can serve as a check on those with the power to allocate resources. Good answers, in turn, will call forth more resources, by empowering the supporters of projects that contribute substantially

to the overall public interest. There is a vast ongoing expansion of data, micro, and macroevidence that can be used to calibrate this analysis and improve it over time.

So we should see this work as a first step and invite those that care about how efficiently global resources are spent in development to reflect on this evidence. We should work to improve the global evidence base and replicate it in different settings, acknowledging that context will matter both for benefits and for costs. And while being impatient for further evidence, we should first and foremost insist on using what is in front of us. We should use this analysis to ask hard questions of those who would propose to spend resources at odds with the best available evidence on likely costs and benefits.

Stefan Dercon and Stephen A. O’Connell