

# Integrating Value for Money and Impact Evaluations

Issues, Institutions, and Opportunities

*Elizabeth D. Brown*

*Jeffery C. Tanner*



**WORLD BANK GROUP**

Independent Evaluation Group

October 2019

## Abstract

This mixed methods study investigates why fewer than one in five impact evaluations integrates a value-for-money analysis of the development intervention being evaluated. This study distills four main insights from combined analysis of 33 semi-structured and unstructured interviews, surveys of 497 policy makers and 16 journal editors, and portfolio analyses of World Bank and worldwide impact evaluations. The study finds that low levels of training in cost data collection and analysis methods, together with a lack of standardization of the value-for-money assumptions (e.g., time horizons, discount rates, and economic or financial cost accounting) limit value-for-money integration into impact evaluations. Further eroding researchers' incentives, demand for cost evidence from the journals that publish impact evaluations is mixed. Ill-defined standards

of rigor undermine editors' capacity to evaluate the quality of value-for-money analysis when it is integrated with impact evaluation evidence. Institutional funders of impact evaluations do not consistently demand that cost analysis be integrated into their funded evaluations. This study finds no evidence in support of the myth that policymakers do not demand cost evidence. Rather, it finds that researchers have few ways of knowing what kind of analysis policymakers need and when they need it. Improving the stock of impact evaluators who are cross trained in value-for-money methods, establishing standards in what constitutes rigor in costing, resolving methodological issues, and improving linkages between policymakers and researchers would lead to greater integration of value-for-money methods in impact evaluations.

---

This paper is a product of the Independent Evaluation Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/prwp>. The authors may be contacted at [jtanner@worldbank.org](mailto:jtanner@worldbank.org).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# **Integrating Value for Money and Impact Evaluations**

## **Issues, Institutions, and Opportunities**

Elizabeth D. Brown  
Jeffery C. Tanner

Key Words: value for money, impact evaluation, cost-effectiveness analysis, cost-benefit analysis, international development  
JEL Codes: 01-Economic Development, 0100 Economic Development: General

# Table of Contents

<b>1. Background .....</b>	<b>1</b>
<b>2. Evaluation Questions and Strategy .....</b>	<b>4</b>
Approach .....	4
Data .....	5
Methods .....	8
<b>3. Findings.....</b>	<b>14</b>
<b>4. Discussion and Conclusion.....</b>	<b>25</b>

## Box

Box 2.1. Defining Value for Money .....	4
---	---

## Figures

Figure 1.1 Impact Evaluations Published Per Year (1990–2015).....	1
Figure 2.1 Composition of Policymaker sample.....	7

## Tables

Table 3.1 Why is VFM so Infrequently Incorporated into Impact Evaluations?.....	15
Table 3.2 Willingness to Pay for Impact, Cost, and VFM Information.....	17

## Abbreviations and Acronyms

3ie	International Initiative for Impact Evaluations
BACO	best available charitable option
CBA	cost-benefit analysis
CBCSE	Center for Benefit-Cost Studies of Education
CEA	cost-effectiveness analysis
CUA	cost utility analysis
DHS	Demographic and Health Surveys
DIME	Development Impact Evaluation
ERR	economic rate of return
GEA	general efficiency analysis
IE	impact evaluation
IEG	Independent Evaluation Group
IER	Impact Evaluation Repository
LSMS	Living Standards Measurement Survey
SIEF	Strategic Impact Evaluation Fund
VFM	Value for Money
DFID	U.K. Department for International Development
USAID	U.S. Agency for International Development
MCC	Millennium Challenge Corporation
NGO	nongovernmental organization
SROI	social return of investment

*All currency amounts are in U.S. dollars unless otherwise indicated.*

## Acknowledgments

The authors wish to thank Patrick McEwan for contributing his advice and knowledge throughout the project's duration, and to Hank Levin, John Strand, and Howard White who provided views and comments on an earlier draft.

Aliza Marcus and Joost de Laat graciously facilitated access to the contact list of Strategic Impact Evaluation Fund (SIEF) subscribers. We are grateful to the 400+ survey respondents from that list. The authors are grateful for feedback from Joy Behrens, Diana Epstein, David Evans, and Emmett Keeler in developing that policymaker survey, and for the assistance of Holly Blagrove in its implementation.

The authors also thank those who met with them for telephone, in-person, or email interviews, including Juan Belt, Logan Brenzel, Annette Brown, Laura Chioda, Joost de Laat, Shanta Devarajan, Markus Goldstein, Penny Hawkins, Sarah Lane, Ariana Legovini, Ruth Levine, Gideon Lukens, Manny Jimenez, Temina Madon, Meghan Mahoney, David McKenzie, Jack Molyneaux, Owen Ozier, Jyotsna Puri, Dan Rosenbaum, Adam Ross, Justin Sandefur, Lyn Squire, Miguel Szekely, Caitlyn Tulloch, Edit Velenyi, Damian Walker, Howard White, and Keith Wood.

The authors also wish to thank editors of the top seven journals that publish impact evaluations in international development who participated in the Journal Editor Survey—*American Economic Journal: Applied Economics*; *Economic Development and Cultural Change*; *Journal of Development Economics*; *Journal of Development Effectiveness*; *World Development*; *Quarterly Journal of Economics*; and *World Bank Economic Review*. Maria MacDicken assisted in implementing the survey of journal editors.

Erik Bloom, Marie Gaarder, Richard Scobey, Mark Sundberg, and Nicholas York provided insights and encouragement throughout. Yunsun Li and Karol Acon Monge provided excellent research assistance. Yezena Yimer gave unflagging administrative and clerical support.

This work was originally used as an unpublished manuscript developed under the auspices of the Independent Evaluation Group (IEG) with financial support of the government of Sweden, for which we are especially grateful.

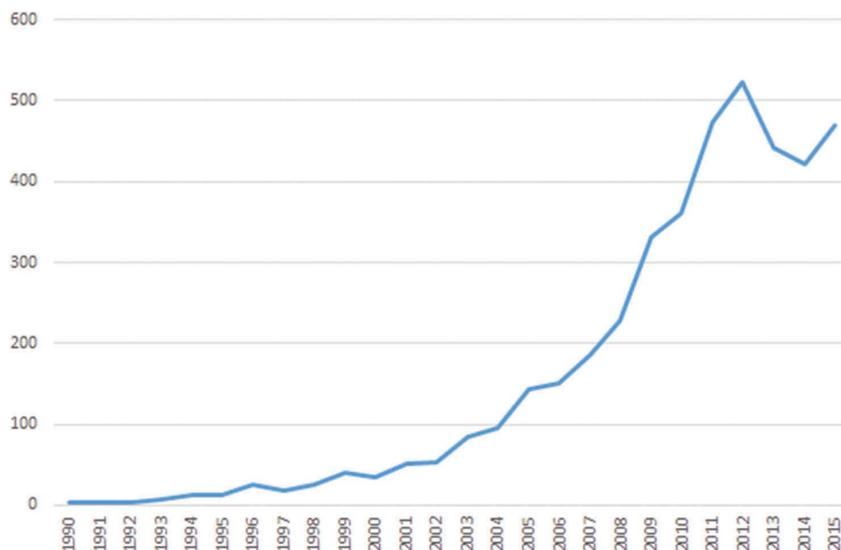
The findings, interpretations, and conclusions are the authors' own and should not be attributed to the World Bank, its Executive Board of Directors, or any of its member countries.

# 1. Background

This paper considers the evidence and incentives for producing impact evaluations that integrate cost<sup>1</sup> data capture and value for money analysis for the purposes of accurately describing the cost and efficiency of an intervention. In this analysis, Value for Money refers to methodologies that measure efficiency, including Cost-Effectiveness Analysis (CEA), Cost-Benefit Analysis (CBA), Cost-Utility Analysis (CUA), Social Return on Investment (SROI), rank correlation, and basic efficiency resource analyses.<sup>2</sup>

For more than a decade, observers have increasingly looked to impact evaluation (IE) methods as a means to evaluate the attributable effects of real-world interventions implemented by development agencies and developing nations (see White and Bamberger, 2008, for example). More recently, however, the confluence of four trends offers unprecedented potential to dramatically increase the production and policy relevance of impact evaluation evidence.

**Figure 1.1 Impact Evaluations Published Per Year (1990–2015)**



Source: Sabet & Brown (2018)

First, the number of IEs on issues important to international development is rising: Over the past 15 years, the number of published international development impact evaluations increased from no more than 50 studies per year before 2000 to between 400 – 500 studies

<sup>1</sup> For the balance of this study, in order to clarify the meaning of the word “cost,” we capitalize Cost when it is used to denote the output of an analytic process as in: *conducting a Cost study* or when referring to a *Cost analysis* or providing policymakers with Cost reporting. We use lowercase when it is used as an input into such an analytic exercise, as in *cost data*, *cost elements*, *costs of activities*, *cost information*. We also use lowercase when cost could be used interchangeably with “price tag” as in *the cost of Cost-related research*, or when it could be used as a verb as in *to cost the financial impact of an intervention*. The distinction can become somewhat nebulous, but it is made in an attempt to clarify concepts and emphasize Cost analysis.

<sup>2</sup> The VFM concept also explicitly considers equity. This project limited the scope of its VFM investigation to the efficiency methods described in this paragraph.

per year between 2013 and 2015, despite a plateau in production after 2012<sup>3</sup> (Sabet and Brown 2018; Cameron, Mishra, and Brown 2015; Savedoff 2013). Second, the internal validity—quality—of impact evaluations also has increased. There are more prospective evaluations of well-defined interventions, more field experiments, and more carefully designed quasi-experiments (IEG 2012). Third, impact evaluations are increasingly asking explicitly comparative questions on relative effectiveness by evaluating multiple treatment arms (Muralidharan and Sundararaman 2011). Finally, as the number of impact evaluations is rising, so too is the number of systematic reviews (White and Masset 2018), which rely on the existence of a range of primary studies in order to uncover generalizable findings.

All four elements—rapid growth in the number of IEs, improvements in internal validity, an increase in IEs with multiple treatment arms allowing for comparability, and an expansion in the number of systematic reviews—create more entry points through which to integrate VFM analysis into project design and decision making.

Yet as the stock of evidence grows, observers have noted that most published evaluations do not contain cost information needed for cost-effectiveness analysis (Dhaliwal et. al. 2013; McEwan 2012) or other analyses of a project’s Value for Money. Institutional production of Value for Money analyses has waned also. For example, the use of economic rates of return (ERR) at the World Bank and other multilateral development agencies has faltered and fallen out of favor. Whereas around 70 percent of the World Bank’s investment projects contained an ERR in the 1970s, about 30 percent did so in the early 2000s (IEG, 2010).

International development scholars whose work integrates impact evaluations with Value for Money analysis have noted the underutilization of CBA, CEA and other efficiency analyses in the published impact evaluation literature (McEwan 2012; Dhaliwal et. al. 2013). However, no analysis to date has determined the percentage of published impact evaluations that incorporates efficiency analysis, and there has been little systematic investigation of the observed decline in VFM application. A 2008 review found that inconsistent use of language; lack of common measures; lack of quality data on social impacts, outcomes, outputs and program cost; lack of incentives for transparency; and the expense of cost measurement all contribute to low levels of Cost reporting in the social sector (Tuan 2008).

Veteran CBA observers note that the low application of VFM analysis in impact evaluation studies and their often-dubious quality has occurred within the context of waning interest in CBA and Cost-efficiency methods more generally. This downward trend is mirrored in the quality of economic analyses in the World Bank’s project appraisal documents. In 2007 and 2008 just over half of the included economic analyses were of “acceptable” or “good

---

<sup>3</sup> Data are from 3ie’s [Impact Evaluation Repository](#). The repository includes published impact evaluations of development interventions carried out in low- and middle-income countries that use experimental or quasi-experimental estimation strategies with a credible counterfactual. 3ie developed the repository using systematic search of over 31 “academic databases in health, economics, public policy and the social sciences provided by platforms such as Ovid, EbscoHost and ProQuest, and libraries and websites from select research organisations and academic institutions.” More than 84,000 potential studies published between 1981 and 2015 were identified and screened, with more than 4,200 (from over 120 countries) selected for inclusion in the repository.

quality”. Identical analysis performed in the 1990s found that roughly 70 percent met the “acceptable” or “good” quality standard (IEG, 2010).

Even so, there are recent examples of influential projects that sought to integrate impact evaluation results with cost-effectiveness analysis to collectively set policy priorities, as with the Disease Control Priorities Project (Jamison et al. 2006). Arguably, as impact evaluation methods continue to mature in sectors like Education, Governance, Agriculture, and Health and Nutrition, evaluators may consider how to re-institutionalize VFM practices.

A major goal of IE studies is to provide scientific evidence on the policies and programs that do and do not work to improve development outcomes. However, if IEs do not include program or policy cost information, resource-constrained policymakers will have limited evidence to guide their selection of efficient programs and policies, or to consider the cost implications of scaling, replicating or reproducing programs and policies found to be effective.

Cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) would be useful for (i) identifying whether social benefits exceed social costs of a single intervention; (ii) comparing worth of interventions with different (monetizable) outcomes; and (iii) comparing the relative worth of interventions that share common outcome(s). CEA and CBA help identify which intervention produces a given amount of outcomes for the least cost. Including VFM analysis may help policymakers understand and use evidence from the growing number of IE field studies in their program decisions by identifying which interventions produce the most outcome for a given cost (White 2014; Dhaliwal et al. 2013; McEwan 2012). Such evidence can provide insight that is “counter to common sense, popular appeal, and traditional ideas” (Levin 2001).

Including VFM directly in IEs yields at least two advantages over producing them separately. First, collecting cost data at the time of the intervention and data collection on the IE outcomes is likely to be less onerous in terms of expense and time than reconstructing program costs ex post. Second, evaluators, policymakers, and other stakeholders can use the evidence to make more efficient decisions when high-quality, rigorous estimates of impact also include an estimate of their Cost.

## 2. Evaluation Questions and Strategy

Despite the potential gains from including VFM analysis in impact evaluations, such integration seems to be the exception rather than the rule. To understand why, we used mixed methods to answer three evaluation questions:

1. How frequently is Value for Money analysis incorporated into published impact evaluations?
2. What are the existing incentives and barriers faced by producers and users, both as individuals and institutions, for VFM incorporation?
3. What are some options to overcome the challenges for the integration of Value for Money analysis into impact evaluations?

### Approach

We investigated these questions from the perspective of those who produce impact evaluation estimates and reports (IE producers), and those who use or who could use those estimates and reports in decision-making for international development policy work (IE consumers). This report defines IE producers as those in the production line of impact evaluations: They pay for, fund, carry out, or communicate evaluations and findings. IE consumers include those who use evaluation results to change behavior, make funding decisions, select programs, or legitimize predetermined behavior. IE consumers include impact evaluation beneficiaries as well as policymakers and decisionmakers.

#### **Box 2.1. Defining Value for Money**

The definition of Value for Money is contested. Our working definition relates specifically to the use of VFM in impact evaluations. In practice, impact evaluations most often use efficiency analyses such as cost-benefit or cost-effectiveness analysis. We adapted a definition from that used by the Department for International Development (DFID) after an extensive review of VFM field methods. DFID defines VFM to include the four “E’s”: measurement of a program’s economy, efficiency, effectiveness, and equity, though in this paper most of the discussion centers around using the quantified attributable benefits identified in impact evaluations as the numerator to establish efficiency with costs as the denominator. The ratio of Impact to Cost yields an estimate of efficiency, recognizing that there are many ways to arrive at estimates of each of those broad constructs.

There are times, of course, when institutions may act as both producers and consumers of both impact evaluations and Cost analyses. For example, a large agency (such as USAID) or multilateral bank (such as the World Bank) could potentially use the evidence from evaluations of vaccine programs that it funded to determine the most effective way of distributing vaccines in resource constrained countries.

We worked through a simple framework of the process by which evaluative evidence is created, disseminated, and applied. That is, we solicited views from impact evaluation

producers, journal editors, and policymakers in addition to estimating the frequency with which efficiency analysis currently appears in published impact evaluations for the first time.

This paper combines results from exploratory research that queried individual impact evaluators, VFM experts, policymakers, as well as representatives of international development multilateral, bilateral, philanthropic, and research institutions and academic peer-reviewed journals using semi-structured and unstructured interviews, surveys, a small randomized control trial, and portfolio analyses of World Bank and worldwide impact evaluations.

## **Data**

Data for the evaluation were procured through five data collection activities:

**NEAR CENSUS OF WORLD BANK IMPACT EVALUATIONS** – A data set of 30 World Bank impact evaluations with VFM analyses was constructed from a sample of 168 World Bank impact evaluations produced between 2000 and June 2010.<sup>4</sup> Of the 168 impact evaluations, 40 were found to contain “a simple comparison of costs with benefits, cost-benefits analysis, economic rate of return, or cost-effectiveness analysis across treatment types or programs” (IEG 2012). Three were dropped because they were falsely tagged as including efficiency analysis when they did not, and two because they did not report the information needed to make a determination. Five IE publications could not be found. The remaining 30 World Bank IEs with VFM analysis were reviewed in detail.

**SAMPLE OF GLOBAL IMPACT EVALUATIONS FROM THE DATABASE OF THE INTERNATIONAL INITIATIVE FOR IMPACT EVALUATION** – We obtained a random sample of 236 impact evaluations from the 3ie’s [Impact Evaluation Repository](#) (IER), representing 10 percent of IER studies that were published between 1986 and 2012. The IER records all published impact evaluations of the effectiveness of development interventions that were identified through systematic search of over 30 databases, search engines and websites. Included interventions must have been carried out in low- or middle-income countries using recognized experimental or quasi-experimental estimation strategies.

**SEMI-STRUCTURED AND UNSTRUCTURED INTERVIEWS WITH IE PRODUCERS AND CONSUMERS, INDIVIDUAL RESEARCHERS, INSTITUTIONAL REPRESENTATIVES, AND POLICY MAKERS** – Thirty-three face-to-face or telephone interviews were conducted with a purposive sample of impact evaluators, impact evaluation funders, policymakers, and individuals specializing in cost-benefit and cost-effectiveness methodologies to learn about the range of institutional barriers and incentives of integrating Value for Money analysis into impact evaluations.

**SURVEY OF POLICYMAKERS**—In partnership with the World Bank’s Strategic Impact Evaluation Fund (SIEF), we designed a survey that was emailed to 3,623 individuals on a contact list maintained by SIEF of policy-minded individuals. These individuals were asked

---

<sup>4</sup> A list of the World Bank impact evaluations that were included in the sample is available from the authors.

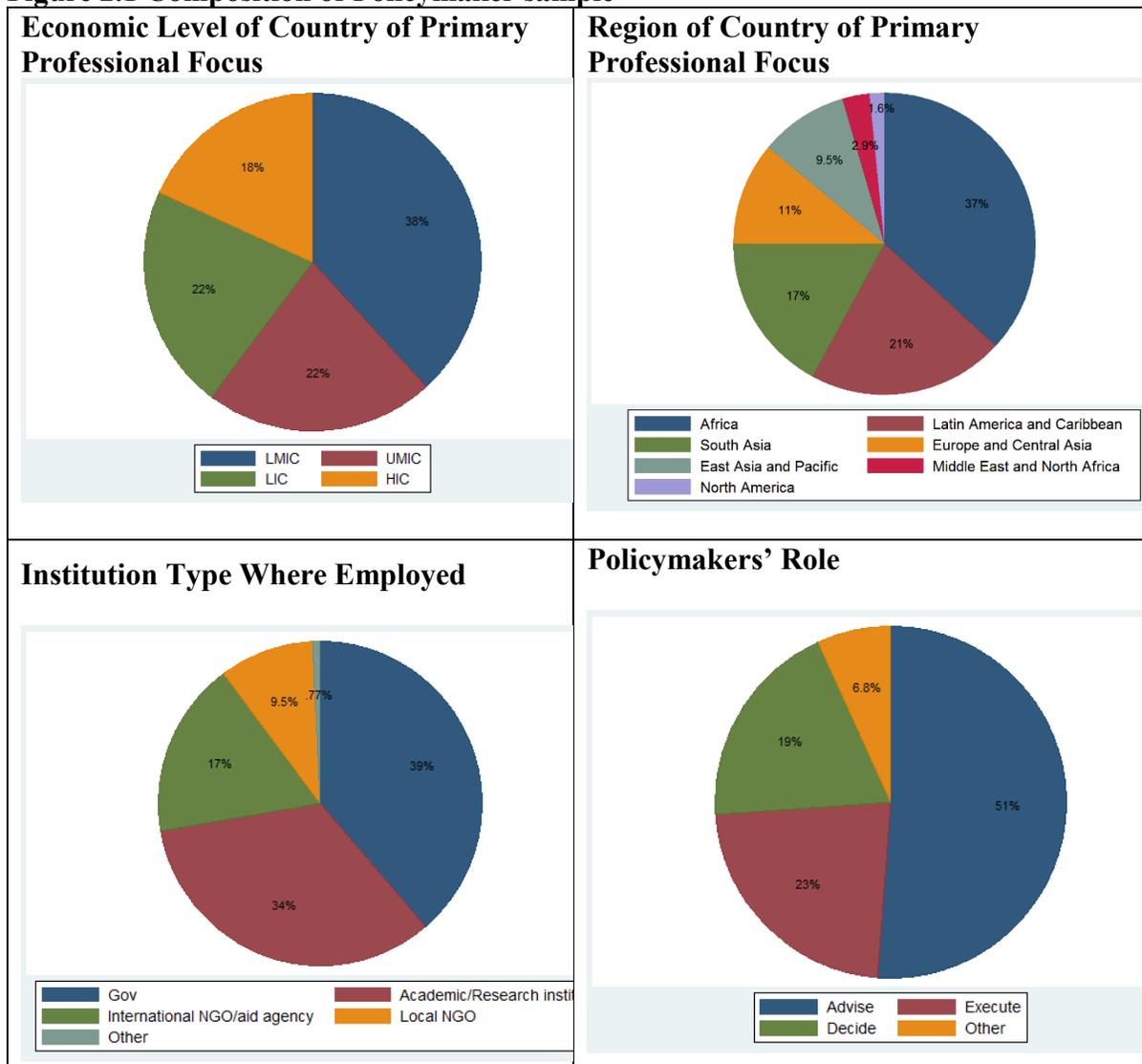
## EVALUATION QUESTIONS AND STRATEGY

to respond to five questions related to Value for Money to understand their willingness to pay for Impact, Cost, and VFM studies.

The survey had 497 respondents who answered at least one question; the analytic sample is composed of 407 individuals who answered at least one of the “willingness to pay” questions. The survey asked respondents about their role within their organizations and then mapped those roles onto three functions within the policy-making process: Advisors (composed of researchers, academics, evaluators and consultants), Decisionmakers (high level executives within government or international or local NGS), and those who Execute those decisions (government functionaries, project managers); a category for “Other” (e.g. teachers, students, or librarians).

Of the analytic sample, two-thirds were male. Respondents were well educated. Nearly one-quarter had a Ph.D., almost two-thirds had a master’s degree, 1 in 10 had a bachelor’s degree, and less than 1 percent had anything below a bachelor’s degree. They also tended to be well established in their careers: 60 percent were between 35 and 55 years old; 17 percent were older than that; and 23 percent were between 25 and 34. The income levels and regions of the country they professionally work on are shown in figure 2.1. So, too are the types of institutions where they work and their roles within that institution.

**Figure 2.1 Composition of Policymaker sample**



Source: IEG, SIEF policy-maker survey fielded in December 2015.

Note: LMIC: lower-middle income country; UMIC: upper-middle income country; LIC: lower-income country; HIC: high-income country. Specifics on World Bank country classifications can be found at <http://data.worldbank.org/about/country-and-lending-groups>

Country income levels are rather evenly represented, with more than 4 of 5 respondents working in or on developing countries. Respondents are most likely to work in a government post or in academia. Half of the respondent population identifies as having an advisory role, more than 40 percent report a policy role in making or executing decisions. Overall the respondents provide a well-rounded sample of the many stages and places of policymaking.

The 14 percent response rate denotes a self-selected group of those interested in responding to a request for a “short 7-10 minute survey to help the World Bank learn how people like you think about and use the evaluations that SIEF and other groups at the World Bank

## EVALUATION QUESTIONS AND STRATEGY

produce.” The SIEF list itself is a self-selected group of individuals involved in the policy-making process that have self-identified as having an interest in impact evaluations.

**JOURNAL SUBMISSION REQUIREMENTS AND SURVEY OF JOURNAL EDITORS**—We collected the stated submission requirements from the websites of seven journals: The American Economic Journal: Applied Economics; Economic Development and Cultural Change; The Journal of Development Economics; The Journal of Development Effectiveness; World Development; The Quarterly Journal of Economics; and The World Bank Economic Review, and responses to an email survey of the journal’s editorial staff.

The stated submission requirements describe each journal’s expected standards of rigor, topics of interest to its readership, and whether the journal accepts empirical or theoretical subjects. None revealed any preference for a specific methodological approach, including approaches such as CBA, CEA, or other VFM methods. Given the absence of formal guidance on VFM methods, we developed a five-question survey to explore journal editor’s perspectives and opinions and the journals’ de facto policies and practices.

1. Do the journal’s stated editorial policies or submission requirements discuss cost-effectiveness, cost-benefit, and Value for Money analysis?
2. What formal and informal editorial practices govern the inclusion of cost-effectiveness, cost-benefit, and Value for Money analysis in published impact evaluations? And, what are they?
3. Do submissions with cost-effectiveness, cost-benefit, and Value for Money analysis receive special consideration?
4. Why is VFM so infrequently incorporated into impact evaluations?
5. What opinions do journal editors have regarding whether published impact evaluations should or should not include VFM analyses?

The survey was sent to 44 editors by email in December 2015. Sixteen journal editors responded to the survey, a 36 percent response rate representing six of the seven journals contacted.

## Methods

The team employed a mix of qualitative and quantitative methods to answer the research questions. We applied:

- Qualitative methods to separately analyze responses from the 33 structured and semi-structured interviews; the 16 responses to the journal editor survey; and the set of 30 World Bank impact evaluations that contained a value for money analysis.
- Statistical methods to separately analyze the sample of 236 impact evaluations drawn from 3ie’s repository; the 30 World Bank impact evaluations that contained a value for money analysis; and the 497 responses to the policymaker survey.

## ASSESSMENT OF WORLD BANK IMPACT EVALUATIONS CONTAINING VFM ANALYSES

We developed a framework to analyze the frequency, type and transparency of VFM analyses found in the sample of 168 IEs produced by the World Bank. Information from the subset of 30 IE's found to contain any kind of VFM IE's was tabulated in a Microsoft Excel spreadsheet for analysis by Sector Board, business line, method, and transparency of reporting.

The analysis generated three outcome variables corresponding to the type of VFM analysis observed in the studies: CBA, CEA, or general efficiency discussion (includes some VFM, but no indicators meeting the CBA or CEA criteria, and no VFM (a small number of tagged studies were found to contain no VFM analysis).

An impact evaluation was classified as having a cost-benefit analysis (CBA) if it included a comparison of estimates of Costs and Benefits (with data of costs and benefits) and one or several of the following indicators: Benefit to Cost Ratio, Cost to Benefit Ratio, Economic Rate of Return/Internal Rate of Return, Financial Internal Rate of Return, Present Value, or Net Present Value. The studies classified as CBA vary in the extent of their reporting from quick, one paragraph "back of the envelope" calculations, to more elaborate exercises. For example, we classified as a CBA the analysis in "Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach" implemented in Bolivia. This IE includes benefit-cost ratios in the range of 2.28 and 3.66 resulting from sensitivity analysis of the program's benefits under different assumptions. We classified "Does Management Matter? Evidence from India" as a "back of the envelope" CBA. Firms included in the study did not want to report internal accounts information. The evaluation's authors reported a 130 percent rate of return in one year based on an analysis of the firm's main cost (cost of the consultancy firm's services) and an estimate of other costs. In a third example, we classified as CBA "Contracting-Out Dialysis in Romania: What Was the Impact?" which compares the total program Cost with the total savings of the benefitting public entity.

An impact evaluation was classified as having a cost-effectiveness analysis when it compared two or more alternative programs measuring the same outcome. Cost comparisons to programs with different outcomes were excluded from CEA classification. We classified as CEA two Cambodian experiments that presented the cost per unit increase in student promotion, dropout, and achievement in literacy and numeracy to the same impacts in a control group. We classified less direct comparisons as CEA if a comparison was present. For example, one CEA compared the marginal cost per child of an intervention that provided information to parents in Pakistan to the cost of programs with the same outcome in low-income countries.

The general efficiency analysis (GEA) designation was applied when the IE included an efficiency discussion that did not qualify as either CBA or CEA or as another VFM method.<sup>5</sup> We classified IEs that reported average intervention costs, cost per beneficiary, the share of

---

<sup>5</sup> Cost utility analysis, social return on investment, basic efficiency analysis, multi-criteria appraisal, rank correlation of cost versus impact, cost minimization analysis.

## EVALUATION QUESTIONS AND STRATEGY

administrative cost, and other non-comparative discussions of Costs and Benefits as GEA. For example, we classified as GEA “Seaweed Farming in Indonesia” based, in part, on the following statement: “SEApIant’s cost has been high relative to benefits derived to-date. IFC committed roughly \$1.9 million to the SEApIant program through 30 June 2006, and this investment has not yet yielded significant tangible returns in terms of either increased earnings for farmers or the introduction of value-added processing.”

We documented the cost indicators and ingredients by type, e.g., opportunity costs, discount rates, timeframes, inflation, and exchange rates between local currency and the currency used in the analysis, for each IE reviewed in detail. Transparency was evaluated along six dimensions: 1) methods clarity (i.e. statement of a specific, identifiable method), 2) cost reporting (i.e. listing of ingredients and their value), 3) analysis (i.e. level of aggregation, unit cost per impact, unit cost per input, treatment of opportunity cost), 4) Cost adjustment reporting (i.e. reporting discount rates where appropriate, inflation adjustments, and currency exchange rates), 5) rationale (i.e. explicit rationale for a VFM analysis), and 6) study limitations, (i.e. reporting partial analyses such as the exclusion of costs related to some benefits derived from the program, or the use of approximate costs, data quality issues, missing data, or focusing on only a short period for accounting benefits).

## FREQUENCY OF VFM IN GLOBAL IMPACT EVALUATIONS

To estimate the frequency of VFM analysis in global impact evaluations, we first developed a set of keywords to represent each of the VFM methods based on extensive review of the VFM literature. Keyword sets were developed to represent cost-effectiveness analysis, cost-benefit analysis, cost-utility analysis, financial analysis, social return on investment, basic efficiency resource analysis, multi-criteria appraisal, and rank correlation of cost vs impact.

A text analytics algorithm was conducted over a full-text search of each of the 236 impact evaluation studies to generate tables of keyword frequencies. This method also was applied to a “benchmark sample” of 65 studies drawn from a database of 168 World Bank Group impact evaluations studies published between 2002 and 2011 (World Bank Group impact evaluation database).

The benchmark sample was formed by combining IEs from two groups of World Bank impact evaluations. The first group contains 35 IEs identified as including some kind of (unverified) efficiency analysis. The second set of studies was pulled randomly from the remaining (untagged) 131 studies contained in the near census of World Bank impact evaluations described above. The outcome variable for the untagged impact evaluations is “No VFM.”

Since IEs in the first and second group do not have the same probability of being selected into the benchmark sample, we applied a sampling weight. The probability of selection into the benchmark sample is 1 for the 35 tagged studies. The probability of selection for the untagged studies is  $30/131 = 0.229$  for the 30 untagged studies. Hence the sampling weight,

calculated by taking the inverse of the selection probability, is different for tagged (weight = 1.057) and untagged studies (weight = 4.37).

To estimate the (unknown) proportion of impact evaluations in the 3ie sample that contains any kind of efficiency analysis, we first ran an ordered logistic regression on the benchmark sample with the sampling weights applied, where the categorical outcome variable is defined as:

$$VFM = \begin{cases} 3 & \text{if formal Vfm analysis} \\ 2 & \text{if general discussion of Vfm} \\ 1 & \text{if No VFM analysis} \end{cases}$$

The explanatory variables originate from keyword frequencies indicative of presence of (formal) VFM analysis. We grouped the keywords into three mutually exclusive categories: (i) return on investment/cost-benefit analysis, (ii) cost-effectiveness, and (iii) decision analysis of other kinds. Next, we constructed three explanatory variables<sup>6</sup> as the total frequency of the words mapped to each category.

We predicted the outcome for the 3ie IEs using the estimated parameters from the benchmark regression and explanatory variables constructed in the same way for 3ie studies. This requires the assumption that the relationship between the VFM outcomes and key word frequencies in the World Bank's IE portfolio also holds in the 3ie sample. The probability of having any kind of VFM analysis in the 3ie sample is  $1 - Pr(VFM=1)$ .

#### **PRODUCER AND CONSUMER PERSPECTIVES ON VFM IN IE**

The team analyzed responses from eight initial unstructured interviews to first identify key topics related to VFM in IE that were used to inform the design of the semi-structured interview protocol. The semi-structured interview protocol covered: (i) the interviewee's perspective on the importance of VFM in IE, (ii) their institution's formal or informal rules governing the production and use of VFM in IEs, (iii) sources of demand for VFM in IE, and (iv) actions that the interviewee thought would increase the production or consumption of VFM in IEs.

All interview responses were recorded using notes taken during the interview and transcribed into electronic format for analysis. A single analyst reviewed and coded responses according to emergent themes in the data. Analysis explored the multiple reasons respondents gave in response to a question by enumerating unique responses and tabulating response frequencies. Because the statistical representativeness of the key informant sample is unknown, we do not engage in statistical analysis for this sample.

---

<sup>6</sup> Alternative constructions of explanatory variables include: (i) using total frequencies of all words/terms as the single regressor; (ii) grouping the words/terms into six groups, giving rise to six instead of three explanatory variables; and (iii) getting regressors from principal component analysis. Regression and prediction results are qualitatively similar. The three variable model has the advantage of being parsimonious while also distinguishing between formal VFM methods (e.g. CEA and CBA) and informal VFM discussions.

## **EVALUATION QUESTIONS AND STRATEGY**

Rather, this purposive sample was constructed to parsimoniously cover the range of roles of those engaged in the production chain of IEs. Ten of the selected interviewees worked at multilateral agencies, seven worked at international NGOs, six worked at bilateral agencies, four were from large foundations active in international development, and two apiece worked in a research institute or the Executive Office of the President. Finally, IEG interviewed one individual currently working in an academic institution.

The sample included eight individuals whose primary role was to design and develop analytical tools, guidelines, and templates for economic efficiency analysis such as cost-benefit, cost-effectiveness, impact modeling, and Costing studies. Ten interviewees were primarily involved in designing and carrying out impact evaluations at a research institute, a multilateral agency or a non-governmental organization.

Ten interviewees primarily served in the capacity of a funder. Their decision-making often included determining the basis for awarding grants to impact evaluation teams, monitoring progress on the evaluations, reviewing impact evaluation results, and disseminating findings.

Five interviewees function primarily as policymakers. Two policymakers provide training, assistance, review, and support to agencies in fulfilling government-wide policies related to efficiency analysis and evidence standards. Two worked in large, highly influential research institutes or NGOs that help to define international standards of evidence for impact evaluations in international development.

Many of the individuals interviewed had a secondary role of influencing or setting their own institution's policies. For example, all the funders interviewed also consume and use the results of the impact evaluations and Cost analyses to inform grant-making decisions and institutional policies. Likewise, many impact evaluators use evaluation results to recommend policies and programs, or to identify areas for further and future research.

## **POLICYMAKER'S WILLINGNESS TO PAY FOR COST AND IMPACT INFORMATION**

Our online survey of more than 400 individuals involved in the policymaking process in developing nations presented a hypothetical scenario in which an unspecified education program had generated test score improvements of 0–20 percent in other countries. Respondents were told that they had an opportunity to implement the program for up to 10,000 students in their country, but that they could also choose to do an impact evaluation to learn of the program's true effects in their setting.

Respondents were asked how many of the 10,000 students they would be willing to have not receive the program in order to be able to pay for the impact evaluation. A similar question was then asked about the number of beneficiaries they would forgo in order to do a Costing study.

Out of concern for framing effects, the order of the first and second of these questions—on impacts and Costs alone—was randomized. Simple t-tests revealed that order did not have an effect on responses for willingness to pay for either Effectiveness or Cost studies.

A third question asked about respondents' willingness to pay (WTP)—again in terms of excluded students—to do a study that would tell them efficiency by giving them both the Costs and the benefits. Here we did find evidence of framing effects. Those who see the impact question first are willing to pay 629 (~26 percent) more units for the combination of benefits and Costs as a VFM analysis than those who saw Cost estimates first. All three questions required respondents to enter the number of excluded student beneficiaries as an integer between 0 and 10,000.

### **JOURNAL EDITORS' PERSPECTIVE ON VFM IN IE**

The journal editors' responses to the survey questions were tabulated and enumerated in an electronic spreadsheet. The responses were summarized in tables. Sixteen of 44 journal editors of the seven journals that publish the most impact evaluations in international development responded to the survey, a 36 percent response rate. Qualitative and descriptive analyses of the journal editors' responses are provided in the findings section.

### 3. Findings

Each of the three evaluation questions posed by the report is addressed by triangulating analysis obtained across data collection activities.

#### **QUESTION 1: HOW FREQUENTLY IS VALUE FOR MONEY ANALYSIS INCORPORATED INTO PUBLISHED IMPACT EVALUATIONS?**

We estimate that 18.9 percent of the World Bank’s impact evaluations include any kind of VFM analysis, while the predicted proportion in the 3ie data set is 14.1 percent. Our prediction indicates that the 3ie sample has a somewhat lower, yet not statistically different, proportion of VFM studies compared to the sample of World Bank IEs.

The estimated percentage of IEs with any kind of VFM analysis in the 3ie sample has not changed much over time. The estimated percentage of IEs that include VFM was:

- 16.8 percent of IEs published prior to 2004;
- 11 percent of IEs published between 2005 and 2008; and
- 14.5 percent of IEs published between 2009 and 2012.

A majority of the World Bank’s IEs with efficiency analysis—around 80 percent—conducted a CBA or CEA, and the remaining conducted a general efficiency analysis.

The quality of cost ingredients reporting was mixed in the sample of 30 World Bank IEs. About 13 percent of the CEAs and CBAs listed only the main cost components, whereas just over half listed several cost ingredients; and the remaining third simply reported a Cost estimate without detailing much of what went into it. About 40 percent of World Bank IEs classified as category II (those with any kind of VFM) reported the cost of a program in relation to the number of beneficiaries—a common metric used in VFM analyses.

Fewer than 20 percent of CEA and CBA studies reported on unit cost per impact. These percentages are very low, considering that they are included in impact evaluations where at least one measure of impact is supposed to be very available to be used in these comparisons.

We find that just 25 percent of the CEAs and CBAs reported opportunity costs. However, because policy makers craft decisions based on real budgets and financial costs, excluding opportunity costs may be theoretically imperfect but practically correct.

The transparency of reporting is very low among the 30 World Bank IEs that included any kind of VFM analysis. Descriptions of the Costing data, methods and assumptions are often vague or incomplete. Just 13 percent reported unit costs of inputs, and only 17 percent reported unit costs per impact. Only 8 percent of IE projects conducted over a year reported discount rates and only 21 percent reported adjusting for inflation. There was also a lack of consistency in reporting discount rates, currency conversions, and inflation adjustments. For example, 58 percent of the analyses did not report the exchange rate used to convert from local to the currency of analysis. In addition, very few studies—around 13 percent—provided

an explicit rationale for including Cost Efficiency analysis and just under half explicitly reported the limitations of the Costing exercise. Our analysis suggests the results of available Cost Efficiency analyses from WB impact evaluations should be used carefully. Such analyses often quantify a portion of the program’s benefits, and frequently suffer from missing information and missing cost data.

In sum, our analysis reveals the need to elevate reporting transparency and the need to develop standards of reporting so that the available evidence can accumulate and be used to improve the use of resources in development cooperation. In addition, our analysis suggests the need to examine the root causes underscoring the wide variation in cost-efficiency reporting across evaluators. It is difficult to know why VFM analysis is so infrequent in the absence of further inputs from IE producers and consumers.

**QUESTION 2: WHAT ARE THE EXISTING INCENTIVES AND BARRIERS FACED BY PRODUCERS AND USERS, BOTH AS INDIVIDUALS AND INSTITUTIONS, FOR VFM INCORPORATION?**

The descriptive analytic findings on the state of VFM in IE are symptoms of underlying problems faced in their production, dissemination, and use. We found three primary reasons for the underproduction of VFM in IE—organized around three principal actors: impact evaluators, policymakers, and institutions.

**Finding 1:** The expected payoffs to producing efficiency analyses is dulled by the (inaccurate) perception that policymaker demand for VFM is low when it is more accurately described as uncertain.

As shown in Table 3.1, our sample of IE producers most frequently cite a lack of policymaker demand as the reason for little VFM observed in IEs. IE-producing interviewees supported this argument with three main points: first, interviewees questioned the cost data’s relevance to a policymaker’s individual contextual considerations. Second, many interviewees felt that the political calculus, rather than a project’s investment case, dominates a policymaker’s decision-making. This view of political economy dynamics undercuts incentives to generate efficiency analysis and often detrimentally influences the analysis itself. Third, IE and VFM producers felt uncertain about what policymakers need and want to know and when they need to know it; furthermore, producers felt unsure about how to go about gaining that insight.

**Table 3.1 Why is VFM so Infrequently Incorporated into Impact Evaluations?**

<i>Main reasons</i>	<i>IE Evaluator</i>	<i>IE Funder</i>	<i>CE/CBA Methodologist</i>	<i>Policy-maker</i>	<i>Sum</i>
Policymakers do not demand VFM analysis	6	3	3	3	15
There is little incentive for academic impact evaluators to produce VFM	4	3	3	3	14
Impact evaluators lack consensus on how to apply CEA and CBA methods; there is no agreement on the "right" approach	6	4			10

Carrying out the cost data-collection and analysis is costly to the projects		3	5		8
Obtaining cost data can be challenging	2	4	1		7
Measuring the impact or benefit is the evaluator's first-order concern	2	3	1		6
Large institutions implement VFM in IEs inconsistently across their programs		3	2		5
Institutional conflicts and interests stand in the way of applying VFM to institutional decision-making	2	1	1		4
The skills of VFM and IE experts differ and there is not a lot of integration across projects or within institutions	1	2	1		4
The importance of VFM will increase as more large-scale randomized evaluations are carried out		2			2

Source: Analysis of unstructured and semi-structured interview responses to the question: “Thinking about the field, generally, why do you think VFM is so infrequently incorporated into impact evaluations?”

Even so, producers suggested that the political calculus differed greatly depending on the amount and on the quality of the evidence provided. Validating this point is the fact that impact evaluations have flourished at current levels of policymaker interest, and that quantitative analysis of the policymaker survey data<sup>7</sup> showed that their willingness to pay for causal information on benefits is not much higher than their willingness to pay for Cost information.

Analysis of the policymaker survey of 407 respondents in the study sample revealed that respondents are willing to pay about 10 percent more for an impact evaluation study as they are for a Cost study, but the observed difference is not statistically significant. Likewise, the average willingness to pay for VFM is also higher than willingness to pay to know a projects effects alone, but again this difference is not statistically significant. However, respondents indicated they are willing to pay a 13 percent premium for a study that presents VFM (Costs and Benefits) than for a Cost study alone – a statistically significant difference at the  $p < 0.01$  threshold. Together, these results imply that although decisionmakers are most interested in *what works*, there is also clear demand to know *bang for buck*.

---

<sup>7</sup> Based on a 14 percent response rate of nearly 4,000 individuals on the mailing list of the World Bank’s Strategic Impact Evaluation Fund. These individuals self-selected for the mailing list based on their interest in impact evaluations—which underscores the parity found between willingness to pay for effectiveness information and Cost information. Because a sampling frame of those involved in the policy-making process does not exist, the external validity of these results, as with all findings in this evaluation, are indicative even if not conclusive.

**Table 3.2: Willingness to Pay for Impact, Cost, and VFM Information**

Paired t tests:

WTP	obs	Mean1	Mean2	dif	St_Err	t_value	p_value
IE - Cost	353	2685.2	2428.9	256.2	139.4	1.85	0.067
IE - VFM	344	2674.8	2755.4	-80.6	109.3	-0.75	0.461
Cost - VFM	345	2434.0	2757.2	-323.1	123.9	-2.6	0.009

This exercise abstracts from the true cost of conducting any of these three kinds of study and sets aside the fact that Cost analyses are generally an order of magnitude (or more) cheaper than an Impact Evaluation.<sup>8</sup>

Those involved in the policy process indicate that a study that combines impact and Value for Money should be less costly than the sum of doing those two activities individually. While we find no evidence of survey or framing effects on WTP for Cost or Effectiveness studies, we do find that respondents who were primed by the IE question are willing to pay more for a VFM study than are those who randomly received the question on willingness to pay for Cost first. Those who see Impact first are willing to pay 629 more units (~26 percent) more for VFM than those who saw Cost first.<sup>9</sup>

Respondents indicated that they valued the combination of Cost and benefit analysis more than they valued either Cost or benefits individually. These results hold for nearly every subgroup of functionary in the policy-making process in the survey—advisors, deciders, implementers. These findings undermine the notion often articulated in the interviews that a lack of policymaker demand is responsible for the lack of VFM analysis.

**Finding 2:** Researchers perceive significant cost but little incentive to produce VFM in impact evaluations.

The second most frequent reason for so little VFM is that researchers and evaluators who produce IEs are offered little incentive to integrate Cost analysis. The costs of VFM include uncertainties that affect the time and effort required to carry out rigorous evaluation of policy and program Cost. For example, the lack of consensus on how to apply CEA and CBA methods; the lack of agreement on the "right" approach; and the challenges of obtaining Cost data in the first place and effort to analyze the data all disincentivize VFM production.

---

<sup>8</sup> Our own experience and discussions with Cost analysis experts and IE producers put the price of a Cost study at \$20,000-\$40,000, while the price tag for an impact evaluation typically runs from \$400,000-\$2.5M.

<sup>9</sup> Perhaps priming with the higher cost IE item first leads to a proclivity to spend more for combined products. It would be useful to repeat this experiment to understand with greater certainty whether the observed framing effect is merely a statistical artifact (e.g. a Type I error).

One of the driving incentives for IE producers—whether in academia or (to a lesser degree) in international organizations—is the opportunity to publish in professional journals. From a researcher’s perspective, there is little incentive to take time out of the preferred activity—measuring benefits—to collect high quality cost data and work through tricky cost estimates and assumptions to achieve a VFM analysis with the same level of rigor as the IE if it is not rewarded.

Evidence from our journal editor survey indicates that demand for VFM from top journals is mixed at best. Perhaps most illustrative of the low esteem held for VFM by journals is the fact that while nearly all of the top economics and economic development journals that we investigated had statements on the standards of estimating benefits (often termed effects, outcomes, results, and so on), none addressed quality control criteria for Cost estimates, much less efficiency. Additionally, 13 of 14 journal editors surveyed reported neither formal nor informal editorial practices governing inclusion of (CEA, CBA, or other VFM analysis) in published impact evaluations. Twelve of the journal editors did not give special consideration to VFM analysis—an IE with VFM is generally no more likely to be published in the top academic journals than is an IE without VFM analysis.

Journal editors mainly thought that Value for Money is so infrequently incorporated into impact evaluations because so few studies can obtain high quality cost data. “While such calculations are helpful in ‘ball parking’ the magnitude of an intervention, and whether it is ‘worth it,’ I suspect most researchers don’t believe that the estimated numbers are of sufficient precision or reliability to require people to put dollar figures on interventions.”

Several journal editors noted methodological issues and a lack of standardization could cause the analysis to be rejected: “Because careful valuation of Costs and benefits involves difficult decisions on certain parameters, such as shadow prices, discount rates and welfare weights, which may be arguable, and thus expose authors to additional rejection risks from referees.”

Two editors said they will ask authors to remove VFM analyses. Two others noted that CBA, if done carefully, could stand on its own and deserved to—rather than be crowded in with impact estimates.

Still, about half of the editors indicated that VFM should be included in IEs. Collectively the Journal Editors provided three stipulations for including VFM analysis: each case would need to be considered on its own merits; the analysis would have to be done well and convincingly; and the journal’s quality standards would have to be met. Editors indicated that VFM analysis with a high level of rigor could be reported independently, although not necessarily in a top journal—the currency by which many IE producers exercise influence and secure tenure.

Indeed, although journal editors raised the notion that CBA, if done properly, can potentially stand on its own, there are few if any examples of this in those editors’ own journals. Scrutiny for VFM quality in journals is de facto left to peer reviewers, who are almost certain to be inconsistent in the standards that they apply. That inconsistency adds another layer of uncertainty—and so risk—to a prospective author.

Even if coordination of methodological standards was resolved, challenges to external validity remain. As pointed out by Evans and Popova (2016), program Costs differ greatly across contexts even for the same type of intervention. These Cost differences naturally result in large variations in efficiency estimates. Uncertainty in effect estimates, recall bias in expenditures, and scale can substantially influence cost-effectiveness estimates (Evans and Popova 2016). Even so, transparency in Cost reporting and assumptions can improve cross-context comparisons and usefully quantify the nature of the variation in implementation costs. Better still, multi-arm evaluations within the same context can largely evade such thorny issues.

Impact evaluators in general seem hungry for templates, checklists and third parties to coordinate and guide their efforts. Such tools have the potential to effectively lower the effort required to build and justify new models and assuage the coordination concerns of evaluators. However, there is a lack of harmonization between the institutions that commonly utilize such tools, leaving still unanswered the coordination failure of having consistent and comparable methods.

Even so, if developed with a core consensus, a revised toolkit has the potential to crystalize accepted standards. To address the concerns about generalizability and sensitivity to specifications, these tools will likely need to be focused on a particular sector (or even intervention or outcome); allow users to explicitly model uncertainty; choose between reasonable methodological alternatives; and to transparently adjust the parameter values of the model's assumptions to more closely align with the particulars of a specified target context and scale.<sup>10</sup> Until such models are available, systematic reviews and league tables such as those done at JPAL<sup>11</sup> and in the innovative work in the state of Washington, USA<sup>12</sup> can provide policymakers with useful information on relative cost effectiveness.

Beyond methodological considerations, data access and quality problems were also cited as significant barriers. Evaluators frequently cite difficulty in extracting reliable expense information—even from administrative data—to reflect the financial cost considerations most relevant to policymakers. Though hardly limited to the World Bank, low levels of baseline financial data collection at the World Bank impedes the ability to conduct CBA at either the start or the end of projects (IEG 2010). Economic costs are even more challenging. Data on opportunity costs are rare, notwithstanding their centrality to economic analysis. Even more challenging than these considerations of basic cost data are the important issues of apportioning Cost (and often benefit) components when an evaluated project is part of a larger multi-arm effort with multiple outcomes and benefits.

These data and methodological challenges reflect an apparent lack of up-front planning to do VFM types of analysis. While the measurement of benefits is now approached in impact evaluations with careful planning well before implementation begins, it is unusual for Cost

---

<sup>10</sup> The second paper in the VFM Component 2 series is a case study that pioneered a tool to address many of these concerns while generating locally relevant, user-informed valuations of the environmental effects of World Bank projects across the globe.

<sup>11</sup> <http://www.povertyactionlab.org/policy-lessons/education/student-participation>.

<sup>12</sup> <http://www.wsipp.wa.gov/BenefitCost>.

analysis to receive the same forethought. Impact evaluators frequently face a lack of cost data and lack of standards for performing analyses. Moreover, some economists—the discipline that churns out most of the IE work in international development—relate that Cost analysis does not “feel like” microeconomics, and together with a lack of training in Cost analysis methods and the challenge of getting it right even if the data are available, make the cost of doing the analysis greater than the benefit—from the researcher’s perspective (Evans, 2016). Together, these factors conspire to expose researchers to risk of costly critique from peer reviewers—convincing many to avoid the exercise altogether.

When it is performed, VFM analysis often seems to come as an afterthought, with cost data inquiries sometimes made well after a project has been closed. This is reflected in the general lack of detail on VFM analysis in impact evaluations, including those done at the World Bank and on Bank projects where cost data should be easier to procure, and Cost analysis should be more valued than analysis done as academic exercises.

For example, of the less than 19 percent of World Bank impact evaluations that include any type of Value for Money analysis, only 13 percent—or less than 2.5 percent of World Bank impact evaluations—reported unit costs of inputs. When cost data are not explicitly available, they are estimated—sometimes borrowing from other studies— but more frequently without meaningful details of how Costs were estimated. Such ex post data collection is prone to recall bias and underestimating expenditures (Evans and Popova 2016). About two-fifths of the World Bank’s impact evaluations do not indicate which cost elements are included in the analysis. Apart from the very real challenges of using accurate cost data, it is extremely rare for World Bank IE-VFM studies to indicate parameters used for basic assumptions in Cost analysis: discount rates, exchange rates, or time horizons, for example. In general, there seems to be a clear disconnect between the careful planning of the estimation of benefits through impact evaluations and the suboptimal quality in the reporting of Cost information. This may explain the hesitancy of journal editors to include it (or institutional funders of IEs require it) as a matter of course.

The combination of these weak incentives and relatively high costs likely accounts for much of the reason why the Value for Money of development interventions is calculated using impact evaluation estimates so infrequently.

**Finding 3:** At the institutional level, political considerations and significant heterogeneity in VFM approaches and methods are constraints to greater inclusion of VFM in impact evaluations.

Individual evaluators and institutional representatives largely agreed that most of the demand for VFM is internal to institutions that both produce and use impact evaluations. However, even for those institutions, there are real challenges to embedding Value for Money analysis within impact evaluations. Political interests and inconsistency in the application of guidelines and methods are clear friction points.

Veteran CEA and CBA methodologists pointed to institutional interests that work against the strict application of cost-effectiveness thresholds to internal decision-making about programs. While the World Bank has largely diluted Value for Money considerations from

its funding decisions (IEG 2010), institutions with clear guidance on integrating Value for Money as a key aspect of their decision process—such as the Millennium Challenge Corporation (MCC)—face difficult internal discussions on whether to fund (or continue funding) a project when an economic rate of return does not meet the institution’s threshold. All institutions interviewed that had a mandate of doing Value for Money assessments reported that VFM analysts often face pressure to find positive results or are at risk of being ignored in decision-making when they do not. Politicized decision-making tends to yield politicized CBA. Many institutions struggle with how to publicize ex post VFM results based on impact evaluations that do not yield the hoped-for results.

Variation between and within institutions of how Value for Money methods are applied contributes to the coordination failure undermining production of VFM in impact evaluation. The challenges of integrating the work of VFM and IE methodologies typically arise in two forms—consistent application of guidelines across groups within an organization and fostering active lines of communication between IE and VFM specialists—across three classes of producer-user institutions: large development institutions (such as multilateral development banks), bilateral and executive agencies (such as the Department for International Development [DFID], U.S. Agency for International Development [USAID], and MCC), and nongovernmental organization (NGO) impact evaluator/funders (such as the Gates or Hewlett Foundations, JPAL, or 3ie).

Larger development institutions tend to struggle more with the challenges of consistent guidelines and active communication between VFM and IE specialists. Although the large development institutions (especially the World Bank and the Inter-American Development Bank) have codified guidance on the inclusion of efficiency analysis throughout various stages of reporting,<sup>13</sup> they often fail to generate consistent guidance about how to integrate VFM in impact evaluations. As a result, VFM can become prioritized by some groups and deemphasized by others within the same institution. For example, even though World Bank leadership has called for an increase of CBA in its impact evaluations, there is considerable variation in the take-up of that directive among the several IE hubs at the World Bank. One World Bank IE hub has implemented explicit requirements that all impact evaluations funded through the hub must also implement a CBA. To aid in this, that hub has collaborated with other highly regarded organizations to generate a [new note](#) to assist in capturing cost data for rigorous Cost analyses. In contrast, another impact evaluation hub at the World Bank has cited a lack of budget and staff expertise in its decision to not incorporate VFM analysis in its impact evaluation work, despite being endowed by a large trust fund from a donor keenly interested in VFM. In general, larger institutions tend to struggle with facilitating collaboration between impact evaluators and those who develop VFM analyses; indeed, these groups are often located in very different departments and rarely communicate.

The bilateral and executive agencies that we interviewed were selected based on having an institutional track record of doing both VFM and Impact Evaluations. Although each has

---

<sup>13</sup> For example, the World Bank’s Project Appraisal Reports (OPSPQ 2013a), Implementation Completion Reports (elaborated when a project is closed), and Project Evaluation (OPCS 2006) reports all give guidance on efficiency analysis. Moreover, IEG’s guidelines for post-completion assessments, Project Performance Assessment Reports (PPARs), also include guidance for evaluating the use of efficiency analysis in the project (IEG 2013). Finally, the WB also has an Economic Analysis Guidance Note (OPSPQ 2013b) that provides more specific guidelines on this issue.

institutionalized VFM within their projects and programs, or are subject to regulations that demand VFM, general VFM policies do not necessarily translate into more VFM analysis in impact evaluations in practice. Each of the government agencies that has such an institutional requirement governs evaluation using an “arm’s length” relationship to retain independence. As a result, agencies have decentralized the decision about whether to carry out an impact evaluation on any given program, resulting in uneven production of VFM based on IE results within any given institution. The MCC tends to do better than most in overcoming the VFM-IE integration challenge. Roughly half of MCC’s projects are impact evaluated. Although these evaluations are always done by third-party contractors, MCC economists are engaged in the design and monitoring of both the intervention and the impact evaluation and are responsible for developing an Economic Rate of Return model before a project is approved and then revisiting that model using the effectiveness results and cost data gathered from the impact evaluation.

VFM generally appears more frequently in NGO impact evaluation producers. These institutions are perhaps the most progressive with respect to integrating and institutionalizing VFM into impact evaluations. Representatives reported formal, institutional practices and guidance governing the inclusion of VFM. Even so, not every impact evaluation included a VFM analysis, either because this guidance applies to certain pools of funds, or because it was difficult to enforce the guidance under all the scenarios in which the grants were made.

### **QUESTION 3: WHAT ARE SOME OPTIONS TO OVERCOME THE CHALLENGES FOR THE INTEGRATION OF VALUE FOR MONEY ANALYSIS INTO IMPACT EVALUATIONS?**

Respondents’ proposals for increasing VFM in impact evaluations ran along three main channels, directed at the entire ecosystem of VFM production and consumption.

1. The most prevalent response to the question of how to increase the production of Value for Money analysis in impact evaluations was the need to develop closer ties to policymakers in order to understand their demands for information. Respondents proposed more efforts to increase demand for VFM from policymakers through improved understanding of their political pressures and needs for research and information and the general space in which they make decisions. Such outreach could take place through workshops to consider how VFM can inform evidence-based policy decisions or could use other strategies described in Dhaliwal and Tulloch (2012).
2. Lowering VFM research costs was another top priority. Respondents proposed (i) investing in and promoting standardized methods (including by information and training on the existing standards); (ii) refining methods which resolve challenges often faced by impact evaluators; (iii) organizing effective peer review by donors; (iv) organizing existing research (through systematic reviews or in league tables); (v) promoting “operationally relevant” VFM analysis in IE performed in more policy-oriented settings such as at the World Bank, USAID, and DFID among others; and (vi) promoting the creation of interactive efficiency models and tools as research becomes more plentiful around a particular sector; this can also go some distance to

assuage a degree of external validity concerns by allowing changes to assumption parameters to fit a targeted intervention context.

3. Communicating findings and facilitating discussion and agreement on methodological issues was a final priority IE-oriented NGOs and communities of practice within and across IE funders can play a significant role in expanding the ‘market of ideas’ around VFM-IE issues. Academic journals can also contribute by publishing clear guidelines on accepted standards for Value for Money analysis to which their peer reviewers will be expected to hold fast, and publication outlets can signal the improved likelihood of being accepted for publication that VFM analysis brings to an impact evaluation.

These proposals highlight and attempt to address the challenges of integrating Value for Money analysis into impact evaluations. Other solutions likely exist. The goal of this paper is to motivate further dialogue among policymakers, impact evaluators, funders, and journal editors to overcome the existing structural barriers and weak incentives and to produce more and higher quality work on the important topic of efficiency. If achieved, the integration of VFM and IEs can make significant contributions to guiding local and global policy decisions on selectively investing in international development interventions, leading to more rapid reductions in poverty, improved economic growth and human welfare. In short, development practitioners will be able to do more, faster, with existing development budgets.



## 4. Discussion and Conclusion

This paper explored the challenge and potential solutions to integrating Value for Money analysis into impact evaluations. It found that current levels of integration are low. We estimate that fewer than one in five impact evaluations have any type of Value for Money assessment.

Several formidable challenges account for this low level of production. Impact evaluators believe that demand from policymakers is low; evidence presented here, however, reveals no statistically significant difference between policymakers' willingness to pay for IE evidence and willingness to pay for Cost evidence—despite the fact that IEs are more than 10 times more expensive. This implies that Cost analyses and VFM exercises represent normative bargains in the eyes of policymakers. Yet IE funders' demand for VFM in IE is mixed and is often driven by the preferences of their donors—donors' reporting requirements for VFM appear to be inconsistent across funding windows and recipients. Academic researchers do not have significant incentives to produce VFM in IE as neither publication outlets nor other professional considerations give significant additional weight for including VFM; on the other hand, academics face non-negligible disincentives. Finally, VFM analysis receives little forethought at the survey design phase, and data collection on costs is expensive and less precise when done after a project closes.

Perhaps most important, there is a lack of cohesive guidance and tools of VFM methods in their application and acceptance among impact evaluators—whose skillsets are generally thin on VFM methods to start with. Templates can help, but specialized expertise is needed to appropriately adapt such templates to the evaluation research design.

When an institution does have formal guidelines that cover efficiency analysis in the formulation of IEs, the application of those guidelines is often uneven, and decisions on whether impact evaluations receive a Value for Money analysis are often uncoordinated at the project level. VFM analysis can increase the value of an IE as a public good. And as with most public goods, provision is sub-optimal when decisions are made individually and in isolation, and when there is no recourse for internalizing externalities.

In addition to low levels of inclusion, VFM analyses in IEs often lack transparency and comparability. Opaque reporting of methods, data sources, prices and discount rates make it difficult for consumers to establish comparability and accurately interpret the results of CE and CBA analyses performed in different settings and countries.

Under increasing pressure to deliver to the last mile, and improve results while reducing costs in its applied policy work, several reviewers of this paper noted that the World Bank has a unique opportunity to advance VFM in IE practice using the following instruments:

- Training Workshops for Staff and Country Representatives, sustained through continuing guidance and technical assistance;<sup>14</sup>
- Cost platforms or templates, together with training on those platforms as needed, to assist users in the appropriate application of VFM methods;
- Greater dissemination and promotion of studies that use CEA and CBA methods so that evaluators and authorities become accustomed to their methods, findings, and implications;
- Codification (or even more pointed guidance) of a common set of methods that is sufficiently flexible to allow for reasonable alternatives that can be specified with a discussion of a basis for deviations from the “standard” and the probable consequences. A standard set of defaults could also be used where the user is not sure of the best assumptions;
- Bolstered vigilance on the quality of efficiency calculations in Project Appraisal Documents and Implementation Completion Reports;
- Expectations of sensitivity analysis on deviations from standard methods and assumptions; and
- Consensus on standard outcome measures and specifications for their construction. These may come from measures derived from the Demographic and Health Surveys (DHS), and the Living Standards Measurement Survey (LSMS), for example. Relatedly, the field needs policy-relevant standard numerators for CEA in different sectors, as has been done with the disability adjusted life year (DALY) in health, for example, or months of additional schooling in the Education sector.

If some evaluators opine that demand for VFM from policymakers may not appear to be terribly strong, it is useful to recall that initially, impact evaluations were not in high demand by policymakers either. Instead, the demand for impact evaluation largely came from development funding agencies and actors interested in basic research. The former have become increasingly interested in VFM in more recent times, in part as a result of domestic resource constraints leading to increased scrutiny of the value achieved from the aid budget. As for the academic world, the demand for VFM as exemplified through its coverage in academic journals is still decidedly less strong than for impact evaluations. There is reason to believe that this could change. One of the appeals of VFM is the ability to compare multiple intervention options. As there are more impact evaluations generated around a theme, efficacy comparisons become more feasible—through systematic reviews and other vehicles. Subsequently, the academic conversation around impact evaluations will likely turn to two frontiers: generalizability and efficiency.

---

<sup>14</sup> For example, a training program sponsored by the Center for Benefit Cost Study in Education, funded by the Institute of Education Science in the U.S. Department of Education includes five days of intensive, hands-on training in cost analysis with technical training in shadow pricing and sensitivity analysis, all based around an open-source platform for performing cost analysis called CostOut (Teachers College, Columbia University 2016).

In the near term, the generalizability question can be partially resolved by examining the similarities in contextual factors between the intervention evaluated by an IE and the specific context where a replication of that intervention is being considered. However, this hinges on the transparent reporting of relevant contextual factors. As pointed out by Waddington et al. (2012), the context specificity of a single study is a strength in generating locally relevant insights, and a weakness when looking to draw more generalizable conclusions. In aggregate, there is a generalizable robustness for interventions that have consistently demonstrated a meaningful effect across multiple contexts. Systematic reviews can help. For a given topic, they accumulate all available evidence that passes risk of bias assessments and present the current state of knowledge on the efficacy of an intervention or interventions, and the number of systematic reviews is increasing.<sup>15</sup>

On the efficiency frontier, the question of Value for Money is ripe for examination, especially because clear identification of effects is a necessary condition to being able to ascribe efficiency. In his 2001 treatment of a very similar question—the lack of Cost Effectiveness analysis in education—Cost Effectiveness veteran Hank Levin drew on 30 years of experience to give three possible explanations: a lack of training, a lack of effects, and a lack of demand by policymakers (Levin 2001). Since that time, the proliferation of impact evaluations and systematic reviews has made large strides on the issue of a lack of effects. And as indicated by the research in the present study, policymakers appear to be nearly as concerned about Costs as they are about effects. The component of production with the greatest gap is capacity of evaluators to do VFM work. If this can be resolved through improved training, Value for Money may see expanded implementation and integration in impact evaluations.

The challenges and barriers to greater production and use of VFM analyses in impact evaluations chronicled in this paper can be overcome by adopting approaches outlined to give greater voice to policymakers' needs, increasing calls and use for VFM from development agencies, increasing incentives for impact evaluators by (inter alia) lowering expected costs, and reducing frictions between supply and demand for incorporating policy-relevant Value for Money analysis into impact evaluations.

---

<sup>15</sup> For an excellent primer in how to do a systematic review, see Waddington et al. (2012). To find systematic reviews, see the Campbell Collaboration ([http://www.campbellcollaboration.org/international\\_development/index.php](http://www.campbellcollaboration.org/international_development/index.php)), Cochrane Collaboration (<http://www.cochrane.org/>), 3ie (<http://www.3ieimpact.org/en/evidence/systematic-reviews/>), and IEG (<http://ieg.worldbank.org>).



## 5. References

- Acumen Fund. 2007. *The Best Available Charitable Option*. New York: Acumen Fund.
- Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja. 2014. "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets." mimeo.
- Bjorkman, Martina, and Jakob Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *Quarterly Journal of Economics* 124 (2): 735–769.
- Boardman, Anthony E., Wendy Mallery, and Aidan Vining. 1994. "Learning from EX Ante/ Ex Post Cost-Benefit Comparisons: The Coquihalla Highway Example." *Socio-Economic Planning Sciences* 28 (2): 69–84.
- Cameron, Drew B., Anjini Mishra, and Annette N. Brown. 2015. "The Growth of Impact Evaluation for International Development: How Much Have We Learned?" *Journal of Development Effectiveness* April, 1–21. doi:10.1080/19439342.2015.1034156.
- Cerdan-Infantes, Pedro, and Christel Vermeersch. 2013. "More Time Is Better: An Evaluation of the Full-Time School Program in Uruguay." The World Bank's Gender Impact Evaluation Database, Washington, DC.  
<http://documents.worldbank.org/curated/en/2013/08/18329366/more-time-better-evaluation-full-time-school-program-uruguay>.
- DFID (U.K. Department for International Development). 2011. *DFID's Approach to Value for Money (VfM)*. London: DFID.
- Dhaliwal, I., E. Duflo, R. Glennerster, and C. Tulloch. 2012. "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education." In Paul Glewwe, ed., *Education Policy in Developing Countries*. University of Chicago Press. Also, <http://www.povertyactionlab.org/publication/cost-effectiveness>.
- Dhaliwal, I., E., and C. Tulloch. 2012. "From Research to Policy: Using Evidence from Impact Evaluations to Inform Development Policy." *Journal of Development Effectiveness* 4(4):515-536. DOI: 10.1080/19439342.2012.716857. Also, <http://www.povertyactionlab.org/publication/research-policy>
- Evans, David K., and Anna Popova. "Cost-Effectiveness Analysis in Development: Accounting for Local Costs and Noisy Impacts." *World Development* 77 (2016): 262–276.
- Evans, David. "Why Don't Economists Do Cost Analysis in Their Impact Evaluations?" *Development Impact* (blog post) <https://blogs.worldbank.org/impactevaluations/why-don-t-economists-do-cost-analysis-their-impact-evaluations>
- Evans, David. "Why Don't Economists Do Cost Analysis in Their Impact Evaluations?" *Development Impact* (blog post) <https://blogs.worldbank.org/impactevaluations/why-don-t-economists-do-cost-analysis-their-impact-evaluations>
- Fleming, Farida. 2013. "Evaluation Methods for Assessing Value for Money. Better Evaluation." Working Group Paper, Australasian Evaluation Society, Perth, Australia.

- Gaarder, Marie M., and Bertha Briceño. 2010. "Institutionalisation of Government Evaluation: Balancing Trade-Offs." Working Paper 8. International Initiative for Impact Evaluation, New Delhi.
- Hewlett Foundation. 2008. *Making Every Dollar Count: How Expected Return Can Transform Philanthropy*. Menlo Park, Calif.: William and Flora Hewlett Foundation.
- IEG (Independent Evaluation Group). 2010. *Cost-Benefit Analysis in World Bank Projects*. Washington, DC: World Bank. [http://ieg.worldbank.org/Data/reports/cba\\_full\\_report1.pdf](http://ieg.worldbank.org/Data/reports/cba_full_report1.pdf)
- IEG (Independent Evaluation Group). 2012. *World Bank Group Impact Evaluations: Relevance and Effectiveness*. Washington, DC: World Bank Group.
- Jamison et al., *Disease Control Priorities in Developing Countries, 2nd Edition Disease Control Priorities*. New York: Oxford University Press; 2006.
- Kremer, Michael, Conner Brannen, and Rachel Glennerster. 2013. "The Challenge of Education and Learning in the Developing World." *Science* 340 (6130): 297–300.
- Levin, Henry M. 2001. "Waiting for Godot: Cost-Effectiveness Analysis in Education." *New Directions for Evaluation* 90.
- Levin, Henry M., Patrick J. McEwan, Clive Belfield, Brooks A. Bowden, and Robert Shand. 2018. *Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis*. Third. SAGE Publications.
- McEwan, P. J. 2012. "Cost-Effectiveness Analysis of Education and Health Interventions in Developing Countries." *Journal of Development Effectiveness* 4 (2): 189–213.
- McEwan, Patrick J. 2012. "Cost-Effectiveness Analysis of Education and Health Interventions in Developing Countries." *Journal of Development Effectiveness* 4 (2): 189–213.
- Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119 (1): 39–77.
- New Economics Foundation. 2004. *Social Return on Investment: Valuing What Matters*. London: New Economics Foundation.
- Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–249.
- OPCS (Operations Policy and Country Service). 2006. *Implementation Completion and Results Report Guidelines*. Washington, DC: World Bank.
- Operations Policy and Quality Department (OPSPQ). 2013a. *Investment Project Financing-Preparing the Project Appraisal Document (PAD)*. Washington, DC: World Bank.
- Sabet, Shayda Mae, and Annette N. Brown. 2018. "Is Impact Evaluation Still on the Rise? The New Trends in 2010–2015." *Journal of Development Effectiveness* 10 (3): 291–304. <https://doi.org/10.1080/19439342.2018.1483414>
- Savedoff, W. 2013. "Impact Evaluation?: Where Have We Been? Where Are We Going?" Presentation at the CGD-3ie Conference, Center for Global Development, July 17, Washington DC. <http://www.cgdev.org/sites/default/files/Savedoff.pdf>
- Stokey, Edith, and Richard Zeckhauser. 1978. *A Primer for Policy Analysis*. New York: W.W. Norton & Company.
- Teachers College, Columbia University. "CBCSE Methods Training." Center for Benefit Cost Studies of Education. <http://cbcse.org/> (accessed 8 September 2016).

- Tuan, Melinda T. 2008. "Measuring and/or Estimating Social Value: Insights into Eight Integrated Cost Approaches." Prepared for Bill & Melinda Gates Foundation, Washington, DC.
- Waddington, Hugh, Howard White, Birte Snilstveit, Jorge Garcia Hombrados, Martina Vojtkova, Philip Davies, Ami Bhavsar, John Eyers, Tracey Perez Koehlmoos, Mark Petticrew, Jeffrey C. Valentine, and Peter Tugwell. 2012. "How to Do a Good Systematic Review of Effects in International Development: A Tool Kit." *Journal of Development Effectiveness* 4 (3): 359–387.
- Weyrauch, Vanesa, and Gala Diaz Langou. April 2011. "Sound Expectations: From Impact Evaluations to Policy Change." Working Paper 12, International Initiative for Impact Evaluation.
- White, Howard. 2014. "Current Challenges in Impact Evaluation." *European Journal of Development Research* 26 (1): 18–30.
- White, Howard and Michael Bamberger. 2008. "Introduction: Impact Evaluation in Official Development Agencies." *IDS Bulletin* 39 (1): 1-11.
- White, Howard, and Edoardo Masset. 2018. "The Rise of Impact Evaluations and Challenges Which CEDIL Is to Address." *Journal of Development Effectiveness* 10 (4): 393–99.  
<https://doi.org/10.1080/19439342.2018.1539387>.
- IEG (Independent Evaluation Group). 2012. *World Bank Group Impact Evaluations: Relevance and Effectiveness*. Washington, DC: World Bank.  
[http://ieg.worldbank.org/Data/reports/impact\\_eval\\_report.pdf](http://ieg.worldbank.org/Data/reports/impact_eval_report.pdf)
- IEG (Independent Evaluation Group). 2013. Independent Commission for Aid. *Guidelines for Reviewing World Bank Implementation Completion and Results Reports: A Manual for Evaluators*. Washington, DC: World Bank.
- Operations Policy and Quality Department (OPSPQ). 2013b. *Investment Project Financing Economic Analysis Guidance Note*. Washington, DC: World Bank.