WORKING PAPER
2016:48

Are there inherent biases in cost-benefit analysis?

Kenneth J. Button and Johan E. Eklund

RESEARCH NETWORK DEBATE
SWEDISH ENTREPRENEURSHIP FORUM

Working Papers Series from Swedish Entrepreneurship Forum

In 2009 Swedish Entrepreneurship Forum started publishing a new series of Working Papers. These are available for download on www.entreprenorskapsforum.se, and are part of our ambition to make quality research available to a wider audience, not only within the academic world.

Scholars from different disciplines are invited to publish academic work with the common denominator that the work has policy relevance within the field of entrepreneurship, innovation and SMEs.

The working papers published in this series have all been discussed at academic seminars at the research institution of the author.

ABOUT SWEDISH ENTREPRENEURSHIP FORUM

Swedish Entrepreneurship Forum is the leading Swedish network organization for generating and transferring policy relevant research in the field of entrepreneurship and small enterprise development.

Swedish Entrepreneurship Forum is a network organization with the aim

- to serve as a bridge between the small business research community and all agents active in development of new and small enterprises.
- to initiate and disseminate research relevant to policy in the fields of entrepreneurship, innovation and SME.
- to offer entrepreneurship researchers a forum for idea sharing, to build national and international networks in the field and to bridge the gap between research and practical application.

Find out more on www.entreprenorskapsforum.se

Are there inherent biases in cost-benefit analysis?

Kenneth J. Button[a,*] and Johan E. Eklund[b]

[a]Schar School of Policy and Government, George Mason University, USA

[b]Swedish Entrepreneurship Forum and Blekinge Institute of Technology, Grevgatan 34, SE-114 53, Stockholm, Sweden

[*]Corresponding author

Email addresses: kbutton@gmu.edu (K.J. Button), johan.eklund@entreprenorskapsforum.se (J.E. Eklund)

Abstract

This note extends the discussion of the potential bias that can exist in cost-benefit analysis. While there is extensive evidence that capture can result in stakeholder manipulation of inputs, there are also claims that the analysis is inherently biased in favor of over-acceptance. The paper shows that, contrary to popular belief, treating projects in isolation is unlikely to produce bias; indeed, it is as likely as not to lead to sub-optimally low acceptance rates.

Keywords: cost-benefit analysis, inherent bias,

Classification codes: D69; D61; L52

Highlights

- Examines the factors leading to up-ward bias in cost-benefit analysis.
- Assesses the importance of using sequential cost-benefit analysis.
- Shows that cost-benefit analysis has no inherent bias.

1 Introduction

A series of empirical studies have found that there is a tendency for cost benefit-analysis (CBA) to overestimate the social gains from many large projects. Physical planners and economists have found that the overestimation of future use and the underestimation of costs are common when performing CBA in such cases (e.g., Button et al., 2010; Flyvbjerg, 2007, 2009). The cause of these systematic biases has largely been attributed to various forms of system capture by vested interests along the lines established in the Public Choice literature by Stigler (1971) and Peltzman (1976).

What has been less frequently discussed is the possibility that CBA is inherently prone to accept too many projects. The works of Flyvbjerg and others have implicitly assumed that the heart of the problem lies in the application of CBA rather than in the intrinsic methodology. We are concerned with the latter and whether those that have focused almost exclusively on system capture have been neglectful of the technical limitations of CBA.

1.1 The theoretical framework

Previous work on this theoretical aspect of the subject is limited, but we concentrate specifically on Hoehn and Randall {henceforth H&R} (1989, 1991), in which they model a single household Arrow-Debreu economy to determine the outcome when conventional cost-benefit techniques are applied across a large part of a public policy agenda. In their analysis, they find that a general error is "… introduced by the independent analysis of a large number of policy proposals put forth by numerous agencies and levels of the public sector". Based on this idea, they argue that conventional cost-benefit techniques fail to correctly evaluate an overall policy agenda because every policy is evaluated independently. The optimal evaluation technique would instead be a sequential evaluation technique in which policy interactions are taken into consideration.

H&R examine two theorems. The first states that if the policy environment is "ε-augmentable", the net benefits will be overestimated, where ε-augmentability means that there are always additional policy proposals with positive net benefits. The second states that, assuming that the marginal cost of policy change is non-trivial, conventional methods will misidentify net benefits as positive when they in fact are negative.

This leads to two conclusions. First, because there is an upper production boundary in an economy, and a conventional CBA has no upper bound, the standard policy evaluation methods is "… certain to overstate a valid measure of net benefits". Second, as the number of policy proposals increases, conventional CBA has allowed too many proposals to pass the test because of non-trivial marginal costs associated with policy change. This leads to the overall conclusion that too many policy proposals pass the cost-benefit evaluation.

If this is so, then the extent of the stakeholder capture of CBA suggested by Flyvbjerg and others is less than claimed. The question then becomes "How general is the model used by H&R?" Much depends on their distinction between the so-called independent valuation and summation (IVS) and the sequenced valuation.

A large part of the problem is of a non-technical nature and concerns the explicit and implicit assumptions that H&R make. Normally, assumptions are either of a standard type or are "more or less plausible" but can also be implausible or even ridiculous. In adopting standard assumptions, it is not always necessary to provide a motivation simply because they are well known among economists. However, when one makes assumptions that are both non-standard and central to the analysis, it is important to make them explicit and provide some reason as to why they are necessary.

In a comment, Quiggin (1991) criticizes H&R, arguing that the model they used is unrelated to the arguments put forward in the article.[1] To some extent, this is true: the problem is that H&R provide few arguments to support the strong and crucial assumptions they make.

In the following, we first take a close look at the nature of policy interactions and then examine the assumptions underlying the two theorems and the conclusions of H&R.

1.2 Independent valuation and summation versus sequential valuation

Two types of cost-benefit techniques are discussed by H&R: the conventional IVS technique and the supposedly optimal sequential valuation technique.

H&R assume that public policy affects a vector of perceived non-market commodities, denoted $\mathbf{s} = (s_1, s_2, \ldots, s_K)$, with the vector being a function of the amount of

[1] A point refuted by Hoehn and Randall (1991).

commodities being produced, $\boldsymbol{\alpha}$. Furthermore, it is assumed that the "… objective of public policy is to regulate the level of $\boldsymbol{\alpha}$ to control the quantity of perceived nonmarket services, $\mathbf{s} = \mathbf{h}(\boldsymbol{\alpha})$".

To examine the outcomes of a conventional cost-benefit analysis and sequential evaluation, H&R distinguish two types of policies. The first is a multi-impact policy where $\mathbf{s}^0$ shifts to $\mathbf{s}^k$. This shift is done with a vector of policy tools, $\boldsymbol{\alpha}^k$, with the post change vector being

$$\mathbf{s}^k = (s_1, \ldots, s_k, s_{k+1}^0, \ldots, s_K^0) \tag{1}$$

Elements s_1^0 through s_k^0 are shifted to s_1 and s_k, whereas s_{k+1}^0 through s_K^0 remain unchanged. The second type of policy only affects one element in $\mathbf{s}^0$, shifting it to $\mathbf{s}_k$. This single impact policy is

$$\mathbf{s}_k = (s_1^0, \ldots, s_{k-1}^0, s_k, s_{k+1}^0, \ldots, s_K^0) \tag{2}$$

H&R assume that a Hicksian compensated measure is applied when evaluating policy proposals. Applying this when using ISV in a policy with g impacts yields the following:

$$ivs(\mathbf{s}^g, \mathbf{s}^0) = \sum_{k=1}^{g} hc(\mathbf{s}_k, \mathbf{s}^0) \tag{3}$$

Therefore, if multi-impact policies are evaluated using IVS, the result will fail to make a correct evaluation. In a policy interaction, the valuation of the *k*th policy proposal is conditional on the *k-1* policy. Applying a Hicksian compensated measure yields the following:

$$hc(\mathbf{s}^g, \mathbf{s}^0) = \sum_{k=1}^{g} hc(\mathbf{s}^k, \mathbf{s}^{k-1}) \tag{4}$$

H&R reasonably find that the IVS technique, if two or more multi-impact policies are proposed, will fail to correctly evaluate the impacts of these policies.

The question is whether H&R's model provides an accurate description of reality. Their analysis is static in the sense that there is an implicit assumption that every policy proposal is put forward and evaluated simultaneously.[2]

[2]The idea that each project in a set of projects may sequentially pass the Hicks-Kaldor test but the outcome shows lower aggregate welfare was initially explored in Gorman (1955) and is different from the idea presented by H&R.

In principle, there are two ways in which policymakers can fail to consider multiple policy impacts or interaction. First, it may be as simple as the cost-benefit evaluations being incomplete, which is very plausible. In this case, a government agency evaluating a *g* impact policy only correctly evaluates *n* impacts, where $n < g$. Second, when two or more agencies evaluate policy proposals simultaneously, they are likely to fail to consider the mutual policy dependence; i.e., agencies A and B independently evaluate their multi-impact policies correctly against a baseline, not recognizing that the policies are mutually dependent. In the latter case, a problem of policy interaction arises, whereas in the former, it is simply a matter of poor cost-benefit methods.

We focus on the second possibility and assume that government agencies correctly identify all *g* impacts of their own policy proposals. The question is then "How severe is the problem of mutual policy dependence?" The fact that H&R implicitly assume that policy proposals are evaluated at a single point in time has implications. In practice, public sector agencies continually revise public policies and re-evaluate policy programs; there is sequencing in public policy making. To model this, $\mathbf{s}_k$ is substituted with $\mathbf{s}^k$ and $\mathbf{s}^0$ with $\mathbf{s}^t$.

In other words, the composition of the vector of non-market commodities, **s**, depends on the point in time at which one looks. This vector is

$$\mathbf{s}^t = (s_1^t, \ldots, s_k^t, s_{k+1}^t, \ldots, s_K^t) \tag{5}$$

where $\mathbf{s}^t$ comprises all public policies prior to time *t*.

Assuming that multi-impact policies are evaluated sequentially over time, the IVS evaluation will involve

$$ivs(\mathbf{s}^{\mathrm{g}}, \mathbf{s}^0) = \sum_{k=1}^{g} hc(\mathbf{s}^k, \mathbf{s}^{t-1}) \tag{6}$$

The conventional IVS technique should therefore yield a result closer to the optimal sequential evaluation technique. H&R (Hoehn and Randall, 1989) write, "Each agency evaluates its proposals as if each were the next marginal increment to the set of baseline policies". However, as public policy can be assumed to be dynamic, the baseline can also be assumed to shift more or less frequently.

If policy proposals are put forward and evaluated over a longer time period, evaluations of the IVS technique approach that of the optimal sequential evaluation technique. This occurs when the baseline shifts between policy evaluations. As these shifts become more frequent, the difference between IVS and the sequential evaluation decreases. Put differently, Equation 6 approaches Equation 4. Given the inevitable time lags in carrying out actions, the IVS is approximately equal to the sequential evaluation technique.

Looking at H&R's first findings, Quiggin (1991) notes the assumption of ε-augmentability being crucial. What he sees H&R saying is that if an economy is producing non-market and market commodities at distance $\bar{a}$ from the efficiency frontier, the economy cannot be improved by more than $\bar{a}$. H&R's argument, however, assumes that the sum of the value of the policy proposals, ε, is greater than $\bar{a}$ or $\sum_{i=1}^{n} \varepsilon_i > \bar{a}$, leading to the conclusion that conventional cost-benefit methods systematically overestimate net benefits as the number of policy proposals increases.

However, Quiggin correctly clarifies that "Theorem 1 of Hoehn and Randall simply states that if there are sufficiently many projects yielding sufficiently large estimated benefits, the total benefits must be overstated. Even for large agendas, it is an empirical question whether the IVS procedure leads to an overestimate of benefits".

H&R also find that when assuming the marginal cost of policy change is non-trivial, standard methods lead to the misidentification of "… the net benefits of both the agenda and at least some of its components as positive when they in fact are negative". While it is true that costs associated with a policy change may be non-trivial, the question is if there is any reason to believe that this is not taken into consideration in conventional policy evaluations. H&R's assumption is nothing more than the idea that the net effect of policy interaction is negative. However, as noted, there is no way of knowing the net effect of policy interaction. Given that interaction effects are both negative and positive, the net effect of policy interaction will approach zero. Hence, conventional evaluation techniques do not result in a systematic bias. H&R's argument only holds if one accepts the assumption that a policy interaction has a negative net effect and that analysts fail to take this into consideration.

To make the point clear, assume that there is no systematic bias in the cost-benefit analysis; a single evaluation may overstate or understate the net benefits, but with many policy proposals, the expect error will approach zero. Now let us assume risk aversion among analysts and policy makers, which is plausible: the result will be that too few proposals pass the cost-benefit test. As Quiggin (1991) noted, only empirical research can resolve these questions.

2 Conclusions

While there is a significant body of work indicating that cost-benefit analysis can be captured in practice by stakeholders — typically politicians seeking to justify an agenda — the issue of whether the underlying methodologies involved are biased toward over-optimism has received much less attention. Hoehn and Randall's innovative work provides a good structure to consider the latter but does not provide strong evidence of systematic bias. Their distinction between IVS and sequential evaluation is illusory, and there is no strong case to support the contention that the effects of policy interaction should be expected to always be positive. They may be equally likely to be negative; indeed, as the number of policies increases, the expected value approaches zero. Thus, there is no definitive evidence that the capture school of thought under- or overestimates the role of governance structures in explaining the ways in which CBA is conducted.

Acknowledgements

Johan Eklund gratefully acknowledge research support from the Marianne and Marcus Wallenberg Foundation.

Declaration of interest

The authors have no conflicts of interest to declare.

Funding

The work of Johan Eklund has been funded by the Marianne and Marcus Wallenberg Foundation.

Role of the funding source

The funding source of Johan Eklund has had no involvement in the writing or submission decision.

References

Button, K.J., Doh, S., Hardy, M.H., Yuan, J., Zhou, X., 2010. The accuracy of transit system ridership forecasts and capital cost estimates. Int. J. Transp. Econ. 37, 155–168.

Flyvbjerg, B., 2007. Cost overruns and demand shortfalls in urban rail and other infrastructure. Transp. Plan. Technol. 30, 9–30.

Flyvbjerg, B., 2009. Survival of the unfittest: why the worst infrastructure gets built—and what we can do about it. Oxf. Rev. Econ. Policy 25, 344–367.

Gorman, W.M., 1955. The intransitivity of certain criteria used in welfare economics. Oxf. Econ. Pap. 7, 25–34.

Hoehn, J.P., Randall, A., 1989. Too many proposals pass the benefit cost test. Am. Econ. Rev. 79, 544–551.

Hoehn, J.P., Randall, A., 1991. Too many proposals pass the benefit-cost test: reply. Am. Econ. Rev. 81, 1450–1452.

Peltzman, S., 1976. Toward a more general theory of regulation. J. Law Econ. 19, 211–240.

Quiggin, J., 1991. Too many proposals pass the benefit-cost test: comment. Am. Econ. Rev. 81, 1446–1449.

Stigler, G.J., 1971. The theory of economic regulation. Bell J. Econ. Manag. Sci., 3–19.

RESEARCH NETWORK DEBATE
SWEDISH ENTREPRENEURSHIP FORUM

WWW.ENTREPRENORSKAPSFORUM.SE