14

Social Welfare Functions in Benefit-Cost Analysis: The ROPI[1] Approach

David Burress[2]

David Burress is a research economist at the Institute for Public Policy and Business Research at the University of Kansas.

[1]ROPI™ refers to a software module now being commercialized as part of KTEC's KTRAK™ project tracking system for technology development. This paper describes part of the earlier research that led to development of the software module.

[2]This work was supported by Kansas Technology Enterprise Corporation and by the University of Kansas. All opinions are the sole responsibility of the author. Helpful comments were received from Pat Oslund and participants at the workshop on Manufacturing Modernization: Learning from Evaluation Practices and Results, Atlanta, September 1996.

Abstract

The sole objective of government in "traditional" benefit-cost analysis is assumed to be maximizing dollars of aggregate generalized income. But empirically, government policy-makers have multiple goals (reflected, for example, in the multiple criteria used in nearly all economic development evaluations); and philosophically, "one dollar one vote" is inconsistent with "one person one vote." This paper describes an approach for determining multiple weighted goals of citizens and introducing them into a formal benefit-cost analysis. This approach was operationalized in the ROPI (Return on Public Investment) model in Kansas. Goals, weights, discount rates, and measurable indicators for the goals were determined with a well-informed citizens' panel, using the "Analytic Hierarchy Process" [Saaty, 1980]. A replication of the panel led to similar results, except that investment-type goals were ranked differently relative to consumption-type goals. The implied social welfare function is a Cobb-Douglas function of aggregative variables that contain non-utility information as well as utility information. The concept of a "benefit-cost" ratio has to be re-examined within this framework. "Benefits" are differences in welfare between the world containing the program under evaluation, and a counterfactual world where program resources are simply wasted. "Costs" are differences in welfare between a world where the resources are handed over to taxpayers, and a world where the resources are wasted. It follows that general equilibrium modeling is essential. The problem of double counting also needs to be reexamined.

Introduction

In benefit-cost analysis (BCA), the traditional approach assumes that government has only one relevant goal: maximizing a social welfare function (SWF) which consists of aggregate generalized income. Most modern benefit-cost theorists reject that assumption [Drèze and Stern, 1987]. In theoretic terms, it is hard to deny the ethical and political importance of distributional goals, and these goals conflict with the "one dollar one vote" assumption.[3] More generally, it is hard to see how "one dollar one vote" could be politically legitimate in a "one person one vote" political system. If not, then "efficiency" needs to be defined in relation to a legitimated SWF rather than the traditional one. Empirical studies uniformly find that real governments do in fact pursue redistributional goals.[4]

Despite these strong objections, BCA practitioners by and large continue to use something close to the traditional objective function. That is true even in the increasingly important case of "social benefit-cost analysis" (SBCA) [Cox, 1986]. In SBCA, practitioners attempt to place dollar values on a variety of non-market and environmental goods, using greatly expanded notions of value including "option value" and "existence value." None the less, the aggregative conception remains "one dollar one vote."

An explanation for this disparity between theory and practice may be that opponents of the traditional approach have been unable to propose and agree upon an operational alternative. A number of textbooks on BCA for underdeveloped countries have proposed using a constant elasticity of substitution function for aggregating individual incomes. However, these texts made no proposal as to the appropriate elasticity of substitution. Another suggestion is that the SWF should be inferred from empirical studies of actual policy outcomes. However, empirical studies have not thus far arrived at a single unified SWF. A third suggestion sometimes referred to as "scoring" is that policy-makers and/or citizens should be interrogated to provide politically-determined relative values to place on non-market goods. However, this approach continues to assume a linear aggregation for market goods that is equivalent to a "one dollar one vote" evaluation of income.[5]

Also important is the fact government goals are multiple rather than unitary, and there are usually no clear political trade-offs between goals. That happens in part because there are many policy-makers engaged in a dynamic political struggle over opposing goals, so that the political equilibrium need have no simple or stable characterization. If there isn't a stable political equilibrium, then there can't be a stable aggregative SWF to describe the policies that

[3]Theorists such as Ng [1984] have tried to defend the traditional approach by arguing that the most efficient government policy would be to maximize the traditional SWF and then handle distributional issues by means of transfer payments. This argument has never been demonstrated to hold in a full general equilibrium model (and almost certainly does not hold in cases where prices change over time.) In any case other theorists such as Haveman [1989] argue it is irrelevant because real government policies do not operate on the Pareto frontier, are not coordinated across policy areas, and are not stable over time. Hence policy should focus on the direction of change resulting from individual policy decisions, and should not focus on the theoretically optimal policy mix.

[4]For citations and discussion on this point, and on other criticisms of traditional BCA raised in this paper, see Burress and Rich [forthcoming].

[5]The ROPI approach as implemented ended up rather similar to the scoring approach, partly because the panel members chose not to place direct distributional weights on income. However, the panel members did include goals such as jobs which are proxies for distributional weights.

result. Policy analysis in such an environment would appear to require multiple SWFs that would separately serve the multiple actors. Empirically, actual evaluations of economic development programs nearly always have included multiple goals -- at the very least, jobs and well as income; often also tax revenues or net fiscal incidence; and sometimes detailed income components such as profits or wages.

At the same time, a case can also be made for providing a unitary or summary SWF that represents an average over SWFs held by individual policy makers or voters. Even at a very low level of political consensus, an average SWF could provide a particular measure of efficiency that can serve as a focus for political discussion. Given a higher degree of political consensus, a unitary SWF could provide clear signals to program administrators about what goals to emphasize in their programs.

This paper describes an approach being used in the Return on Public Investment (ROPI) model in Kansas. The model has been designed for economic development programs in general, but the particular application is to various programs supported by the Kansas Technology Enterprise Corporation (KTEC), including an MEP program (MAMTC, the Mid-America Manufacturing Technology Center). Our approach assumes that there are multiple policy goals, that the nature of these goals should be determined empirically rather than simply assumed, that benefit-cost information should be provided for each goal, and also that some aggregation across goals should be provided.[6]

Ultimate Goals and Intermediate Process Goals

Evaluation methods are driven by the assumed goals of the activity and by the processes encompassed in the activity. It is conventional to distinguish between "outcome" and "process" evaluations. It would be more accurate to say that there is a continuous scale. For example, within the process of technology transfer, there is a hierarchy of outcomes. That is, outcomes of one stage (e.g. patents), are inputs to another stage (e.g. licensing agreements); each stage of the process can be evaluated by looking at its intermediate outcomes. And in the political realm, there is a hierarchy of goals. For example, the program goals of KTEC are related to development of high technology firms in Kansas; but the ultimate political goals have to do with economic goods such as jobs and income for Kansans.

Philosophically, BCA is an "outcome" evaluation. And ROPI as implemented is oriented as far as possible towards the "outcome" end of the scale -- i.e. towards outcomes that reflect the ultimate goals of economic development in Kansas, rather than immediate technology program goals. That is necessary because one general purpose of ROPI is to make fair comparisons between KTEC and other economic development programs in Kansas. However, for practical reasons, some of the indicators used as proxies for these ultimate goals are actually intermediate outcomes rather than ultimate outcomes.

[6]Others at IPPBR who have been heavily involved in the design and implementation of ROPI are Mohamed El-Hodiri and Pat Oslund. Chang-Erh Chou led implementation of the SAM model used for ROPI multipliers. Richard Bendis at KTEC and his predecessor William Brundage have been very supportive of the ROPI project.

Choosing the Evaluation Criteria

But how do we know what the ultimate economic development goals in Kansas actually are? The issue is critical, because it frames the BCA evaluation. Because accurate measurements are expensive and there are a large number of possible goals, it is not possible to evaluate every goal. Instead we must focus on a few, and we need to select the right few. Here are some possible ways to determine the top few goals of a government program:

- rely on the best judgement of the evaluator
- rely on ethical or philosophical arguments
- use the criteria traditionally applied to economic development (income, jobs, and tax revenues)
- use the single criterion used in traditional benefit-cost analysis (generalized total income)
- ask the program administrators
- ask the direct clients of the program
- read the enabling legislation
- ask the legislators who wrote the law, or consult its legislative history
- ask the legislators who are determining the current appropriations
- ask lobbyists and other policy actors associated with the affected stakeholder groups
- ask a "blue ribbon" panel of citizens
- ask the voters what they want from the program using public opinion polls.

There are many serious problems with each approach. The problems generally boil down into trade-offs between accuracy, cost, and legitimacy.

Accuracy

Some approaches are especially subject to potential accusations of bias; for example, academic researchers might be accused of selecting goals based on their own private political agenda; program administrators might be accused of selecting goals which are immediately controllable and easily achieved. Other sources of information are hard to interpret in terms of ultimate goals. The text of the implementing legislation usually focusses on immediate program goals rather than ultimate goals; for example, the legislated goal for KTEC is to

> foster innovation in existing and developing industries, especially the creation, growth and expansion of Kansas enterprises in a diversified range of primary sectors, which develop value-added products, processes and services... [KSA 74-8102].

Blue ribbon panels could be subject to small sample variability, i.e. a different answer might result each time the approach was tried. Public opinion polling has more statistical reliability, but it can lead to uninformed opinions that may not be helpful in technical situations, and uninformed opinion can be especially volatile over time.

Cost

The most informative approaches tend also to be the most costly. Most interview approaches provide poor quality information, unless there is an extensive prior discussion of the conceptual issues involved; but discussion adds to time and cost. Asking legislators to reach agreement takes an exceptional amount of their valuable time, because (by design) legislators are selected from diverse areas, and have many conflicting opinions that are hard to reconcile.

Legitimacy

The approach needs to be politically acceptable, i.e. persuasive within a democratic context. My personal opinion is that the best approach would be for a committee of the legislature to formally determine a set of ultimate goals as evaluation standards for Kansas economic development programs, using a structured study process. Hopefully, they would also meet from time to time and place relative weights on those different goals. It would then be quite clear that KTEC and other agencies were being evaluated against the ultimate goals accepted by those who control the purse strings. This is a group of persons who, in a representative democracy, clearly do have legitimate authority to decide the issue.

Of course, with this set-up the relative weights might change over time when the political equilibrium changes, creating real problems for program administrators. But changing goals is in the very nature of governmental agencies. This applies even to programs with rather clear missions such as education (e.g. should schools teach sex education? civic values?) The mission of an economic development program is relatively less unitary than that of an educational program,[7] and relatively more subject to political negotiation. Explicit legislative input would provide program administrators with accurate and current information about existing political goals for economic development.

But as it happens no legislative committee has determined explicit ultimate goals for economic development in Kansas. Therefore the goals used in the current ROPI model were selected by another approach. In particular, they were selected by an invited panel of citizens from across the state of Kansas (including some legislators) using a process that involved intensive education, discussion, and structured voting over the course of several meetings.[8] We adopted this approach because, in a democracy, citizens are the ultimate source of authority for deciding ultimate goals of government. At the same time, it is expected that decisions in a democracy will be based on informed discussion, and not on superficial opinions from a polling snapshot.

Goals and Weights from the Panels

We repeated the citizen's panel process twice and got reasonably similar results.[9] In both cases we gave the panel an initial list of ten general goals and asked them to refine or modify the list, following several presentations and discussions on economic development in Kansas and on KTEC in particular. The first panel's initial list was complied from a survey of economic development literature. The first panel's final list was used as the initial list for the second panel. In particular, panelists make a pairwise comparison between each pair of goals, weighing the relative importance of a ten percent improvement in each goal. The comparison uses a scale which goes from 1 (no discernible difference), 2 (a barely perceptible difference) through 10 (an absolutely incommensurate

[7] At least as judged by variability among the opinions of members of our two replications of the ROPI citizens' panel, economic development programs have a wide range of ultimate goals, and there is wide disagreement about their relative weight or importance.

[8] Some respected literature on democracy [Dahl, 1989] refers to this general approach as a "policy jury", and recommends that it should be used more widely. However, in a true policy jury panelists would be selected randomly from the general population.

[9] For more complete details, see Burress and Oslund [1994], and Burress, El-Hodiri, and Narayanan [1992]. For an additional application of the ROPI method, see Burress and Oslund [1996, forthcoming].

difference.) According to psychological measurement theory, choices on this scale are likely to reflect ratios of values. A mathematical analysis is then used to find the best representation of these N(N-1)/2 comparisons in terms of N relative weights.

In both panels, the results showed great variation across panel members in the weights that were chosen. Tables 1 and 2 show the goals and the average weights for the two panels. (The goals are summarized by a short phrase in the Tables; goals as stated by the panels involved a somewhat more complex statement.) Average results of the two panels were at least somewhat similar. Income, jobs, and various measures of human capital ranked relatively high. More explicitly distributional goals such as benefits to rural areas or poor people and improving upward mobility ranked relatively low. Non-monetary quality of life ranked rather low. However, the two panels differed with respect to business investment-related goals such as business climate, technology improvement, infrastructure, and entrepreneurial spirit: the first panel ranked them rather low, while the second panel ranked them very high. We might characterize the two panels as being very similar with respect to truly ultimate goals, but differing on the emphasis they placed on consumption now, versus investment now leading to consumption later. In future research, it would be useful to have the panels consider the two class of goals separately (truly ultimate or "consumption"-type goals versus intermediate investment goals) and then place a relative weight on the two classes.

Indicators for Goals

Goals are general and somewhat amorphous statements about what state of affairs is considered desirable. An indicator for a goal is a quantity that can be measured, and that provides some information about how well the goal is being achieved. Because the information in a given indicator is often an imperfect or incomplete reflection of the underlying goal, a particular goal may be assigned multiple indicators

Indicators reflect a compromise between technique and value judgement. On the technical side, evaluation specialists need to determine whether a given indicator can be measured with a reasonable degree of reliability. On the value judgement side, policy-makers or their surrogates need to determine whether the indicator is a reasonable reflection of the desired goal.

In the ROPI model, we selected indicators for the top five goals with the help the second citizen's panel. Results as implemented in the current ROPI model are shown in Table 3. We plan on adding other indicators to the model over time.

Table 1: Results from the First ROPI Panel

Goal		Mean weight	Rank
A.	Jobs	0.1564	2
B.	Income	0.1626	1
C.	Non-monetary quality of life	0.0587	9
D.	Benefiting ordinary people	0.1113	4
E.	Benefiting disadvantaged people	0.0824	5
F.	Benefiting rural areas	0.0683	8
G.	Making physical investments	0.0529	10
H.	Human capital	0.1446	3
I.	Technological innovation	0.0814	6
J.	Entrepreneurial spirit	0.0717	7
K.	Technology Hardware	0.0097	11
	Sum of weights	1.0000	

Table 2: Results from the Second ROPI Panel

Goal		Mean weight	Rank
A.	Jobs	0.1248	3
B.	Income	0.1143	5
C.	Quality of Life	0.0710	7
D.	High Skilled Jobs	0.1516	2
E.	Upward Mobility	0.0572	9
F.	Opportunities in Diverse Regions	0.0705	8
G.	Better Public-Private Organizations	0.0469	10
H.	Highly Skilled Workforce	0.1232	4
I.	Increasing Competitiveness	0.0741	6
J.	Business Climate	0.1663	1
	Sum of weights	1.0000	

Source: IPPBR

.

Table 3: ROPI CRITERIA

Goals	Indicators
1. Good business climate	1. Business start-ups
	2. Businesses that survived 5 years
	3. Federal R&D funds
	4. Non-federal venture capital
	5. Taxes paid by firms
	6. Patents
	7. Physical infrastructure
2. Good jobs	8. Jobs by wage
	9. Jobs by title
	10. Jobs by education
3. Total jobs	11. Total number of jobs
4. Highly skilled workforce	12. Years of experience/education
	13. Workforce training
5. Income and wealth	14. Personal income

Source: IPPBR

The Implied Social Welfare Function

In the ROPI system, each indicator is treated as an individual SWF, and appropriate benefit-cost ratios (BCRs) are calculated. Each indicator is an extensive variable -- i.e. one which can be aggregated by summing across individuals, such as the total number of high skill jobs. (If an indicator were expressed as an intensive variable, such as the percentage of high skill jobs among all jobs, then a BCR would not make sense.) In addition, an over-all BCR is calculated using weighted indicators. The current weights consist of the average goal weights determined by the second panel of citizens. For goals with multiple indicators, the goal weight is prorated across the indicators. The weights are normalized to add up to 1.

Since panelists were asked to compare ten percent increases in the goals, we can interpret the weights to imply that changes in a social welfare index are given by[10]

(1) $dS_t/S_t = \Sigma_i \, w_i dX_{it}/X_{it}$, where

X_{it} = indicator variable
w_i = indicator weight
t = year
i = indicator index
S_t = SWF index.
Therefore the over-all social welfare index we arrive at is

(2) $S_t = \Pi_i \, X_{it}^{wi}$, where $\Sigma_i \, w_i = 1$.

This SWF index itself has units of an extensive variable. Hence BCRs can be calculated for S_t similarly to those for other extensive indicators. Exactly how the BCRs are defined is described below.

[10]Including the factor S_t in the denominator of (1) imposes an arbitrary normalization that doesn't affect the rank ordering expressed by the SWF.

Discounting Over Time

In the ROPI model, as in any BCA, costs and benefits have to be discounted over time. This is true for each indicator, and not merely for income. That raises the usual question of what discount rate to use. It also raises a new question: are the discount rates constant across indicators?

In a traditional BCA, the discount rate for real income depends on the time preference of the representative consumer from whom resources are being withdrawn.[11] The discount rate for other indicators would presumably depend both on consumer time preference and on the rate of change of the real price (or the rate of change of the real shadow price) for the given indicator. In that case, discount rates need not be constant across indicators.

[11]In traditional BCA, it is argued that technical rates of intertemporal substitution, or at least market rates of interest, are more relevant to discounting than discount rates rooted in personal preferences or policy preferences. The reason is that an unconstrained optimizing individual, when saving and consuming over time, adjusts the level of, and hence marginal utility of, consumption at different points in time so as to offset the difference between pure private time preference and the market rate of interest. A similar argument might be applied to discount rates for the various indicators. In particular, the private market economy presumably does provide opportunities for substituting any given indicator in the present for that indicator in the future. However, KTEC cannot easily take advantage of these market opportunities: it generally lacks authority to exceed its annual budget constraint by borrowing, and it faces political pressures against saving ("if you didn't spend it, why did you need it?")

However, Amartya Sen [1982] points out there need be no single representative consumer, and gives reasonable conditions under which the social discount rate can be very different from the private discount rate. In those situations, the discount rate would need to be inferred empirically from policy preferences, using means similar to those discussed above. Again there is no theoretical reason to expect constant discount rates across indicators.

In the ROPI model, discount rates were determined by the second panel of citizens using a structured elicitation procedure. (They selected 12 percent, a rather high rate as compared to rates commonly used in empirical BCA). Various tests detected no significant difference between discount rates for jobs and income. Because of time limitations, we did not explore discount rates for other indicators, but simply assumed discount rates were constant across indicators.

On the Use of Non-utility Information in a SWF

The ROPI SWF index is clearly not a classical Bergson-Samuelson SWF -- it is not expressed as a function of individual utilities; instead it contains much information (such as jobs and patents) that goes beyond utility. Nevertheless, it does provide a well-specified objective function that government can maximize, and hence it does provide a well-specified concept of efficiency.

Does the ROPI index provide some kind of empirical approximation to a Bergson-Samuelson SWF? Certainly not in general -- for example, the expected relationship between an indicator such as new patents and increased utility is likely to be highly variable across states and across time.

On the other hand, S_t arguably does act as a proxy for some Bergson-Samuelson SWF for small changes in a particular economy.[12] But for larger changes in the economy, S_t is likely to be sensitive to non-utility information. In that sense, S is not a true welfare index. In other words, it places values on things in a manner which is partly independent of any direct human preferences for those things.

However, a similar objection can be made to the traditional income-based SWF, or indeed to any other cardinal social choice objective function whatsoever. In abstract economic theory, preferences are simply a rank ordering of outcomes with no cardinal metric attached. Therefore the choice of a cardinal utility metric that can be aggregated across persons always goes beyond strict preference information. In the case of traditional BCA, income (or consumer surplus) is used as an approximate and partial-equilibrium utility index that can be aggregated across persons to form an approximate SWF. The exact general equilibrium SWF to which it corresponds is sometimes called the "money metric", or aggregate equivalent income. This exact SWF consists in the minimum amount of money that would be required at a fixed set of reference prices to purchase sufficient goods to support the given vector of individual utilities. Consequently, the set of reference prices explicitly provides some non-utility information needed to cardinalize the social rank ordering. In the approximate SWF used in traditional BCA, current market prices implicitly provide the reference prices.

[12] In an *ex ante* but perhaps not an *ex post* sense. The general form of the argument is simple: if an indicator is valued by the panel, then there is an expectation on the part of the panel that increases in that indicator will increase some people's utility. Assuming expectations are rational and defining that expectation function formally leads to a (conditional) Bergson-Samuelson SWF. However, we would expect the panel's expectations to change over time due to non-linearities in the economy, if there are large changes in the economy.

In this perspective, the objection many economists make to using non-utility information to support "interpersonal comparisons" seems overstated. Day-to-day politics and policy analysis necessarily abound with cardinal interpersonal comparisons. Sen and other welfare theorists point out other problems in restricting information to pure utility orderings. Arrow's "Impossibility Theorem," for example, is completely dependent on the restriction to pure preference information; under conditions assumed by the Impossibility Theorem, there is no possibility of empirically describing a fair and operational SWF. But BCA does depend on use of an operational SWF. Therefore the whole enterprise of BCA in general, and of this paper in particular, logically implies that some amount of non-utility information must be used in the social objective function.

The money metric does however incorporate non-utility information in more rigorous fashion than does the ROPI SWF. In particular, if the welfare analyst is willing to accept a specific fixed set of references prices as legitimate (which amounts to a rather arbitrary value judgement), then conditionally on those reference prices the money metric is a strict Samuelson-Bergson SWF. Moreover, for small deviations of market prices from the reference prices, consumers surplus has been shown to approximate the money metric. We believe that similar approximations could be established for specific instances of the ROPI SWF along the lines of footnote 11, but this has not been formally demonstrated.

Non-standard Benefit-cost Ratios

We will need some care to clarify the concept of a benefit-cost ratio in the case of generalized indicators of welfare in a general equilibrium setting. In a traditional BCA, the BCR has the form

(3) BCR = B/C, where

B = dollar value of gross benefits of the project, and
C = dollar cost of the project; and hence
B- C = dollar value of the net benefits of the project.

In a partial equilibrium context, all of these concepts seem entirely straightforward. In particular, we ignore the sources of the dollars ("manna from heaven") when we calculate B, and we ignore the uses of the dollars ("sand down a rat-hole") when we calculate C.

But when we extend the traditional BCA into a general equilibrium setting, much more explicitness is required. Net benefits, the simplest case, can be precisely defined as

(4) B- C = SWF(2) - SWF(1)

where SWF(2) is aggregate equivalent income in a world in which the project is done, and SWF(1) is aggregate equivalent income in a world in which the project is not done. In a more general sense, net benefits are just the increase in social welfare from doing the project, as compared with not doing it. (In a retrospective study, SWF(2) would refer to the actual world with the project, and SWF(1) would refer to a counterfactually modeled world without the project. In a prospective study, both worlds would be counterfactual.) To be precise about world "1," we need to specify what happens to the resources freed up by not doing the project; e.g. the dollars might be returned to the taxpayers.

But what exactly are we to make of gross benefits, B? As it turns out, similar precision would require introduction of a third counterfactual world. Exactly which one is not entirely obvious (see footnote 14, and see below for the ROPI approach). C, on the other hand, could be defined in a way that is independent of counterfactuals: it could refer to the change in the budget constraint of the government agency due to the project. Then B could be defined

indirectly using equation (4). That is, net benefits would be defined as the gross change in social welfare, less the change in the budget constraint.

This interpretation makes considerable sense in terms of the optimization problem faced by the government agency, which is

(5) MAX SWF s.t. $\Sigma_i \; C_i \leq C$ (and s.t. actions of all other private and public agents), where maximization takes place over a collection of potential projects, C_i is the cost of the i-th project, C is the total program budget, and i runs over the set of projects that actually will be done. If projects are scalable, it can be shown that the optimum strategy for the agency is to select projects with highest BCRs as interpreted above.[13]

A small problem with this interpretation is that the gross benefits so defined include differential multiplier effects, while the costs do not -- so the BCR is unbalanced and somewhat misleading. In the particular example of a regional model, a project with outside funding can have a BCR>1 and still be unproductive. This problem could be addressed, however, by calculating a threshold BCR based on a null project (for example, a project in which outside funds are handed over to taxpayers in the region). Because of multiplier effects, the BCR of the null project would be greater than 1. A project would then be considered to be productive if and only if its BCR exceeded this threshold value.

A more important problem is that this interpretation can not be extended to an SWF that uses generalized welfare indicators. In such a case, social welfare has different units than budgetary cost -- there is an "apples and oranges" problem. That is, if we define costs in terms of budget dollars, then equation (4) would mix dollar units and welfare units, which makes no sense.

This problem can be solved in two different ways; both approaches are being used in the ROPI model. One approach involves introducing a third modeled world, say "world 3", in which the project resources are somehow wasted (a "sand down a rat-hole" counterfactual). For example, in a regional model, the funds might be contributed to a worthy cause in the rest of the world (ROW). (We must assume there is no feedback to the region from the ROW.)[14] Since ROW persons do not have "standing" to register their utility in a regional SWF, from the point of view of the government objective the worthy cause is a waste of resources. Then we can define costs and benefits as:

(6a) C = SWF(1) - SWF(3)
(6b) B = SWF(2) - SWF(3),

which is consistent with equation (4).

This approach is used for calculating the BCRs in the ROPI model. In particular, opportunity cost for an indicator is defined as the additional amount of that indicator predicted by our economic models if the KTEC budget had been handed over to

[13]In cases where there are discrete projects with fixed sizes, so that there is a problem of exactly using up the budget, selecting projects with highest BCRs sometimes fails to reach the optimum. But even when fitting discrete projects into a fixed budget becomes a factor, the best summary information on the quality of a particular project is still its BCR.

[14]Note that this particular counterfactual would not make sense for a national BCA study. Simply sending dollars abroad without feedback amounts to burning them; burning dollars may have deflationary consequences, but it need not affect real national resource allocations. A "sand down a rathole" counterfactual for a national study would need to assume a physical waste of resources. Exactly how this is modeled affects the measured amounts of B and C. For example, hiring workers to dig holes and refill them is a resource waste with possible Keynesian consequences; while keeping workers unemployed is a resource waste with no Keynesian consequences.

Kansas citizens in the form of a tax reduction (as compared with throwing it away). Because of various multiplier effects, handing new dollars to taxpayers is expected to increase almost any extensive indicator in the region by at least a small amount, so the opportunity cost under this definition is never zero. This general equilibrium concept of cost includes multiplier effects as well as direct budgetary costs, so that a project is viewed as productive if and only if its BCR exceeds 1.

A second approach, also used in the ROPI model, is to calculate a cost-effectiveness ratio instead of a BCR. The cost-effectiveness ratio is defined as the net benefits per budgetary dollar expended. Net benefits are defined as SWF(2)-SWF(1), and budgetary dollars are directly measurable, so the third counterfactual world isn't needed. One advantage of this approach is that cost-effectiveness ratios are relatively intuitive; "net job-years created per million dollars of budget," for example, makes more immediate sense to a legislator than a BCR for jobs. A disadvantage is that cost-effectiveness ratios are not inflation-proof.

The Double-counting Problem

Critics of the scoring method have pointed out the likelihood that the values of social characteristics will be double-counted unless a method is found to prevent it. For example, suppose that a social value is to be placed on "jobs" or "opportunity to work" that is in addition to the income provided by the jobs. (This might reflect the beneficial externalities that result from the dignity and sense of self worth that people receive from their work.) In the absence of special care, estimates of that social value assigned by the scoring method are likely to be unconditional, and therefore contaminated with the social value of the income that jobs will produce, which would already have been counted separately under conventional benefit-cost techniques. Therefore the welfare from an additional job needs to be conceptualized conditionally on a given level of income. That is, we might conceptualize a complex portfolio which adds the job, but then subtracts away an equal amount of income (e.g. with a non-distorting tax.) Critics have doubted whether this kind of conceptualization is feasible for politicians who determine the scores (Lee, 1989).

As suggested above, the ROPI method differs from the scoring method in several respects. Nevertheless, the issue of double counting still needs to be addressed. In particular, we need to ask to what extent the various goals are empirically correlated; and to what extent this correlation affects the relative values placed on goals by panel members. If, as seems likely, goals are valued partly because they correlate with other goals, then weights based on unconditional values will involve some double counting. In particular, the goals that are most highly correlated with other goals are likely to be overweighted through the following type of mechanism:

Assume the true independent weights are α and β for goals A and B; in other words, the considered value judgements of the panel are most accurately described by the SWF:

$$S = \alpha A + \beta B.$$

Assume goal A causes B at an average rate γ.

Assume the panel incorrectly assigns a weight to goal A that partly reflects the causal correlation with goal B; namely $\alpha^* = \alpha + \delta, 0 < \delta$.

Suppose that an increase ΔA in goal A is observed. That increase leads to a correlated increase $\Delta B = \gamma \Delta A$ in goal B. The total value of the change is evaluated by a BCA as

$$S^* = \alpha^* \Delta A + \beta \Delta B.$$

Yet the panel's true political preferences are described by $S = \alpha \Delta A + \beta \Delta B < S^*$.

In the ROPI approach, there are several layers of procedures and assumptions that can address this possibility.

1. The ROPI weighting method is based purely on pair-wise comparisons. Even if there are correlations across goals, this procedure should reduce the first order correlations in the data. That is, the relative ranking of two goals would be influenced only by the difference in their correlations with other goals. If, as seems intuitively plausible, the various economic development goals are generally positively correlated, then this differencing will tend to reduce the effect of any correlations. But of course, some error would still remain in the weights.

2. In our procedure for eliciting value judgements, we asked the panel to conceptualize goals as independent in the sense of ceteris paribus. For example the job goal was to be conceptualized as increases in total jobs while holding total income and all other goals constant. If panelists successfully did so, then the double counting problem has been eliminated.

One could doubt, however, whether the panelists are able to perfectly conceptualize this kind of control, especially when multiple goals are involved. An empirical test would be to ask panelists to rank multi-variate bundles of goods and see if the rankings were consistent with the elicited SWF weights. Given limited time, we did not pursue this idea with our two panels.

3. With sophisticated repeat players, for example legislators on an economic development committee, players would become more aware of the consequences of the evaluation procedure over time, and shift their votes accordingly so as to more truly represent their policy preferences.

4. Even if correlations between goals have contaminated the weights to some extent, the weighted criterion can still be viewed as a welfare proxy, though not as an exact welfare index. In other words, the imperfect index S* is likely to increase whenever the ideal index S increases, although it will not track it exactly.

Similarly, in the construction of time series indices (such as an index of leading indicators) it is common to ignore the correlations between input variables. In the absence of data on their correlations, equal weighting is often used. Provided that all of the input variables are positively correlated with the target variable, this procedure does lead to a valid predictor of the target variable; and the equally weighted predictor is nearly always an improvement over any one input variable.

5. In any case, this double-counting problem applies only to the weighted indicator, and not to the individual indicators. Policy-makers in general are likely to differ on the appropriate weights, and many of them will view the weighted BCR as simply another indicator. Any relatively small errors in the weights may be of limited importance in an actual political context, where the weighted indicator is more likely to serve as a focal point for discussion than as an agreed decision-making tool.

Future Research

It will be apparent that many issues raised in this paper need to be addressed in additional research. Some of the outstanding issues are:

- How consistent are results of different citizens panels?
- How do citizens panels operate empirically? Are there important framing effects resulting from the ways in which issues are presented to the panel? Do a small number of panel members tend to dominate the outcomes?
- How best should we conceptualize and present the structure and hierarchy of goals to the panel? Do panelists agree that distinctions between programs

goals, investment goals, and ultimate consumption goals are meaningful?
- How serious is the double counting problem? How well can it be managed?
- How sensitive are BCRs to choice of variant "sand down the rat-hole" counterfactuals? Can small changes in the counterfactual lead to changes in the ranking of projects?

Conclusion

The ROPI approach is based on a definite sensibility, a set of "meta-normative" assumptions about benefit-cost analysis. In particular we believe:

- Evaluation should be based on multiple goals or value judgements, a multicriterion approach.
- Yet there is also a need for single summative value judgement.
- The sources of the value judgements should be empirical, not based on the views of evaluators.
- Market prices are not an adequate source of information on value judgements. That is, BCA value judgements should reflect the views of policy-makers or voters, not merely the views of those whose preferences are backed with money.
- The procedure for establishing fundamental value judgements that drive the BCA should be consistent with democratic legitimacy.

Some of these assumptions are inconsistent with the traditional approach to BCA. Undoubtedly our approach raises many new research problems, some of which are listed above. At the same time, we believe the ROPI model does demonstrate that an operational alternative to the traditional BCA is a practical endeavor that can be based on existing modeling tools.

References

Burress, David, Mohamed El-Hodiri, and V.K. Narayanan. 1992. An Evaluation Model to Determine the Return on Public Investment (ROPI) for the Kansas Technology Enterprise Corporation. Research Report No. 196. Lawrence, KS: University of Kansas, Institute for Public Policy and Business Research

Burress, David, and Pat Oslund. 1996 (forthcoming). KTEC OUTCOMES, 1987-1995: an Evaluation of the Kansas Technology Enterprise Corporation Using the ROPI Methodology. Research Report. Lawrence, KS: University of Kansas, Institute for Public Policy and Business Research

Burress, David, and Pat Oslund. 1994. Public Returns to Technology Transfer: an Evaluation of the Kansas Technology Enterprise Corporation Using the ROPI Methodology. Research Report No. 212. Lawrence, KS: University of Kansas, Institute for Public Policy and Business Research

Burress, David, and William Rich. Forthcoming. "Economic Analysis and Distributive Justice", Research in Law and Economics

Cox, L.A., Jr. 1986. "The Social Utility Framework for Benefits Assessment -- Theory of Regulatory Benefits Assessment" , Ch. 5, pp. 111-157 in J.D. Bentkover, V.T. Covello, and J. Mumpower, eds., Benefit Assessment: The State of the Art, New York: D. Reidel Publishing

Dahl, Robert A. 1989. Democracy and Its Critics. New Haven: Yale University Press

Drèze, J., and N. Stern. 1987. "The Theory of Cost-Benefit Analysis." In Handbook of

Public Economics, Volume II, Ed. A. J. Auerbach and Martin Feldstein, New York: North Holland

Haveman, Robert. 1989. “Economics and Public Policy: On the Relevance of Conventional Economic Advice” (Thompson Lecture, Midwest Economics Association), Quarterly Review of Economics and Business 29, p. 6

Lee, Douglas B. 1989. “Transit cost and performance measurement”,Transport Reviews 9, pp. 147-170

Ng, Yew-Kwang. 1984. “Quasi-Pareto Social Improvements”, American Economic Review 74(5), December, pp. 1033-1050

Saaty, Thomas L. 1980. The Analytic Hierarchy Process (Planning, Priority Setting, Resource Allocation). New York: McGraw Hill

Sen, Amartya K. 1982. "Approaches to the Choice of Discount Rates for Social Benefit-Cost Analysis", pp. 325-376 in Lind, Robert C., ed. Discounting for Time and Risk in Energy Policy, Baltimore: Johns Hopkins Press.