



Unleashing the Power of Social
Benefit-Cost Analysis:
REMOVING BARRIERS

October 19-20, 2009
Washington Court Hotel
Washington, D.C.

SPONSORED BY

The Benefit-Cost Analysis Center
Evans School of Public Affairs,
University of Washington

FUNDED BY

The John D. and Catherine T.
MacArthur Foundation



Unleashing the Power of Social
Benefit-Cost Analysis:

REMOVING BARRIERS

Welcome from the Benefit-Cost Analysis Center

I would like to take this opportunity to thank Michael Stegman and the John D. and Catherine T. MacArthur Foundation for making this conference possible. This conference is part of a multi-year effort funded by the MacArthur Foundation to improve the standing of benefit-cost analysis and its use for evaluating the effectiveness of social programs.

This effort is only one of the exciting projects going on at the Benefit-Cost Analysis Center, also generously supported by the MacArthur Foundation. The Center supported the founding of the Society for Benefit-Cost Analysis, and has supported the establishment of the upcoming *Journal of Benefit-Cost Analysis*, with Editor Scott Farrow from the University of Maryland—Baltimore County (UMBC). With MacArthur Foundation support, the Center has also undertaken a project to produce a series of white papers leading towards the establishment of principles and standards for use in benefit-cost analysis.

Thank you all for joining us in this second conference exploring the use of benefit-cost analysis for evaluating social programs.

Richard O. Zerbe, Jr.
Director, Benefit-Cost Analysis Center
Daniel J. Evans Distinguished Professor
Evans School of Public Affairs
University of Washington

October 19, 2009
Washington, D.C.





Welcome from The John D. and Catherine T. MacArthur Foundation

Michael Stegman, Director of Policy & Housing

Welcome, everyone, to this conference on improving benefit-cost analysis for social programs. Thank you to Dr. Zerbe, Katie Ward, and Valerie Chang for making this conference possible and for putting together a great program. Before the first panel starts, I want to give a little background about why the MacArthur Foundation is involved in this project.

This conference and the next two—planned for 2010 and 2011—bring together thought leaders in benefit-cost analysis with practitioners, government officials, and policymakers. The intent is for this gathering to discuss the role, potential, challenges, and opportunities of using benefit-cost analysis to strengthen social policymaking.

This project has its origins in the MacArthur Foundation's \$35-million multi-year effort on the Power of Measuring Social Benefits. This effort attempts to put to rigorous test the shared hypothesis that motivates areas of domestic grantmaking at the MacArthur Foundation. We believe that there is a coincidence of interest in people in need or at risk in greater society, and that programs that invest in these people benefit the individuals, benefit the larger economy, and provide tax benefits to society that last long after the program has ended.

The MacArthur Foundation is pursuing a three-part strategy to test this hypothesis. Our first strategy is to expand the number of high-quality, complex benefit-cost analyses of effective social programs. To do this, we support 25 major studies, many piggybacking on evaluations of programs funded by the National Institutes of Health and other government agencies. For example, we are conducting a benefit-cost analysis in partnership with the Robert Wood Johnson Foundation on the Health Insurance Experiment in Oregon, which expands Medicaid programs. We're pursuing studies trying to determine the value-added of programs, measure direct and indirect benefits of social programs, without just focusing on what narrow interests currently exist.

The second strategy the MacArthur Foundation is pursuing is making a commitment to support benefit-cost analysis efforts writ large. This includes supporting the Benefit-Cost Analysis Center at the University of Washington, and working to meet the need for a high-quality





Unleashing the Power of Social
Benefit-Cost Analysis:

REMOVING BARRIERS

publication outlet through supporting the creation of the *Journal of Benefit-Cost Analysis*. We are also committed to supporting the principles and standards effort, while also doing work in shadow-price estimation and other projects that will benefit MacArthur Foundation interests and benefit the broader benefit-cost community.

The third strategy we pursue is supporting organizations that take research and bring it to the policy process. All of these strategies come together in these periodic annual meetings. We welcome all of your participation and look forward to the ideas generated here.

Michael A. Stegman, Ph.D.
Director of Policy and Housing
The John D. and Catherine T. MacArthur Foundation

October 19, 2009
Washington, D.C.





Table of Contents

Welcome from the Benefit-Cost Analysis Center _____	1
Welcome from The John D. and Catherine T. MacArthur Foundation ____	2
Table of Contents _____	4
About the Conference _____	5
Panel I: Barriers to the use of benefit-cost analysis and cost-effectiveness analysis in public health and medical decisionmaking _____	6
Discussion: Barriers to the use of benefit-cost analysis and cost-effectiveness analysis in public health and medical decisionmaking _____	10
Panel II: Barriers to the valuation of health outcomes in benefit-cost analysis and cost-effectiveness analysis _____	13
Discussion: Barriers to the valuation of health outcomes in benefit-cost analysis and cost-effectiveness analysis _____	18
Panel III: Behavioral approaches to policy _____	21
Discussion: Behavioral approaches to policy _____	24
Panel IV: The use of benefit-cost information in education policy—barriers and opportunities _____	26
Discussion: The use of benefit-cost information in education policy _____	32
Panel V: Estimating the willingness to pay of the American public for childhood poverty reduction _____	34
Discussion: Estimating the willingness to pay of the American public for childhood poverty reduction _____	40
Panel VI: Extending shadow prices available for benefit-cost analysis of social programs _____	43
Discussion: Extending shadow prices available for benefit-cost analysis of social programs _____	46
Panel VII: Barriers to policymakers’ use of rigorous research and benefit- cost findings _____	48
Discussion: Barriers to policymakers’ use of rigorous research and benefit-cost findings _____	55
Appendix A: Conference Program _____	56
Appendix B: Conference Participants _____	59





About the Conference

This national policy forum examined the role of benefit-cost analysis (BCA) in social policymaking and highlighted the new opportunities presented by recent uses of benefit-cost analysis in social arenas. The conference also aimed to develop strategies for making benefit-cost analysis more practical, consistent, and implementable within the social policy fields. Leading scholars, practitioners, lawyers, and policy-makers shared ideas about the implementation of benefit-cost analysis techniques and procedures. Conference participants offered advice about strategies to improve the quality of social BCA and increase its use to strengthen social policies and programs at all levels of government.

This summary was prepared by Benefit-Cost Analysis Center staff based on notes taken during the conference. Center staff would like to thank Jason Schwartz for his aid with notes for panel III. This summary is intended to reflect the tenor and salient points of the conference, rather than to serve as a complete transcript of events and proceedings.

The conference was sponsored by the Benefit-Cost Analysis Center at the University of Washington's Evans School of Public Affairs, and funded by The John D. and Catherine T. MacArthur Foundation





Panel I: Barriers to the use of benefit-cost analysis and cost-effectiveness analysis in public health and medical decisionmaking

Setting the stage for the discussion of benefit-cost analysis (BCA) and healthcare, **Lou Garrison** presented **Mark Pauly**'s talk exploring why BCA of medical services is so controversial. To begin, Pauly asserts, health policy abhors BCA in its strictest sense of monetizing benefits, but also in its broader sense. There is considerable political resistance to using any kind of method that considers cost, and there is also internal resistance to using BCA instead of cost-effectiveness analysis (CEA). Figuring out this difference requires both a normative and a positive analysis.

The normative question is: Is there something about medical goods and services that makes BCA suspect, either in absolute terms, or relative to CEA? The positive question is whether, in alternative models of public choice or political economy, BCA as applied to medical goods and services raises issues that are different from the issues raised in other applications?

The obvious normative problem BCA raises is the willingness-to-pay (WTP) problem. Willingness to pay is more of a problem in medical services than in other areas because medical services are excludable private goods, suggesting there is no natural averaging and a greater chance of "bias." Relying on WTP means that health services to high-income people have higher value. Currently, cost analysis is primarily demanded for public programs, Medicaid, and Medicare. For Medicaid, user WTP is low or irrelevant, so why not use CEA with a socially chosen dollar value per quality-adjusted life-year (QALY)? Medicare is primarily insurance for the lower middle class as a promised substitute for private insurance, so preferences matter, but how should they be included? Whether or not private insurance companies use BCA may not matter.

Taking a closer look at Medicare, the valuation should be a blend of public and private WTP, but no one thinks of Medicare that way – it's thought of either as "society's resources," or as a "promise to me." In either of these situations, using either BCA or CEA may be seen as going back on the deal. Possible solutions include convening committees of experts, performing informal CEAs, and allowing people to choose plans that use their preferred costs.

In the private insurance arena, premiums are based on behavior so far, avoiding any

Lou Garrison
University of Washington

John E. Calfee
*American Enterprise
Institute*

Peter Neuman
*Tufts University School of
Medicine*

Panel convened by **Mark
Pauly**, the Wharton
School, University of
Pennsylvania





explicit valuation beyond labeling the “unproven and useless.” Whether or not BCA is adopted could be subject to the market test—though this would need to be thought of in the aggregate, rather than person by person, as England does. The question still remains whether I would prefer the plan that uses BCA—including my WTP—to ration my healthcare.

This scenario brings up the unspeakable problem of rationing. Is there a way to gracefully and rationally ration healthcare? Politics abhors explicit resource allocation based on cost in healthcare, and blames the current high cost of healthcare on greed and waste. Benefit-cost analysis and CEA may be seen as acceptable if they say “yes” to treatment, but if not if they say “no” to something the doctor or patient wants.

Mark Pauly’s suggestions for an argument supporting BCA and CEA in the medical context include showing the benefits and costs in the aggregate; showing the distribution of benefits and costs by income, age, and anything else relevant; and turning the ultimate decision back on the politician. The initial signs are not promising, given the current refusal to consider cost or quality of life, and the political stance against any rationing. However, strong efforts to improve BCA—including measures of WTP and using more persuasive language—would definitely have a high benefit-to-cost ratio. Pauly basically says “Go on, and let’s see what we can make of this.”

John E. Calfee brought the perspective of research in medicine and health to the forum. He noted that the American Enterprise Institute had convened a conference several years ago prompted by the fear that the Federal Drug Administration (FDA) was going to start regulating BCA in pharmaceuticals. This was met with concern that managed care would use cost-effectiveness analysis and other tools as an excuse to cut back on drug use without pharmaceutical companies having the opportunity to present their own view. This did not come to pass, however; managed care does not shy away from drug prescriptions.

Calfee noted that in the context of healthcare practice, CEA has been resisted, and is almost systematically not used in either healthcare practice or healthcare policy. There are reasons to think that BCA will be even more resisted, in part because of the sense that you arrive at an “up/down” determination, and that becomes the decision. Calfee noted that BCA may be more difficult to apply in the healthcare and healthcare technology arenas based on the different product development and marketing schemes, where products are marketed and developed quite intentionally and continuously until the patent expires.





REMOVING BARRIERS

Calfee articulated several unique aspects of BCA as applied to healthcare. On the benefits side, he noted that to a large extent, healthcare marketing is targeting the highest-value uses of a product, and as a result, products often go to populations that have higher-than-average WTP. This leads to greater benefits being created, and also the imbedding of advances in areas that are judged to have greater value. Benefits are hard to estimate in the healthcare context because of continuous research and development. Research and development into making drugs, devices and procedures more efficient continues, and sometimes new drug indications, e.g., second uses for drugs previously approved for other purposes, have more demand than the original use, which can significantly change a previous analysis of benefits.

The costs-side of BCA also poses interesting questions in the healthcare realm. In many cases, research results in decreasing costs for medications and the like, for example, through lower dosages. The logic of clinical trials in order to gain FDA approval means that researchers want to design trials that maximize the chance of getting a positive result. This often results in using a high dosage to gain approval, and after approval, lower doses are tried. These lower doses are often more efficient, less expensive, and have fewer side effects. A tendency toward targeted therapeutics will also bring costs down dramatically.

Even without these challenges BCA will continue to get resistance for the same reasons as CEA does, but also because BCA tends to lead to a decision rule. In an area where the costs and benefits are constantly changing, it may be tricky to use this tool effectively.

Peter Neumann first talked about the growth of CEA in public health in medicine, before transitioning to talk about barriers to cost-effectiveness analysis and cost-benefit analysis, including the preference for cost-effectiveness analysis.

One way Neumann has seen the growth in CEA is through the Tufts CEA Registry (TCR, the Registry), which can help us understand how to better allocate resources. The ability to standardize CEA methods is critical to this effort. The Registry, updated for 2008, includes 2,033 English-language cost-utility analyses (CUAs), 5,313 ratios, and 7,330 utility weights. Based on the studies considered for the Registry, it seems the quality of the field is improving.

Neumann notes there has been an increase in published CUAs, and that the highest intervention type is with pharmaceuticals. Cost-effectiveness analysis has entered mainstream medicine—it is now showing up in the *Journal of the American Medical Association* (JAMA), the *American Journal of Medicine*, the *Annals of Internal Medicine*, and





others. Even though it's often said that the United States doesn't use CEA, 44% of the English-language studies are US-authored.

In addition to the increase in published CUAs, Neumann asserts there has been an increase in quality. Principal investigators are more likely to state the perspective of the study, to state a time horizon, to do incremental analysis correctly, to do a sensitivity analysis and/or a probabilistic sensitivity analysis than 10 years ago.

Neumann hopes that through presenting this information, Medicare and others can be convinced to get more value for their money through using these tools, most immediately through increased use for preventative care. Recent legislation allows for CEA, though it doesn't use the exact words of BCA or CEA.

There are many barriers to increased use of BCA and CEA in health, Neumann argues. These include methodology, bias (e.g., through industry funding), a sense that the "wrong" things are being studied, the decentralization of the U.S. health system, and the political and cultural resistance to rationing. In terms of methodology, there is some sense that the field is not quite ready for widespread use and it isn't based on time-trials or randomized trials. Differences in standards and methodology contribute to this non-acceptance.

The perception of bias through industry funding may be overstated, but a publication bias does exist. The focus on drugs is also a challenge, Neumann asserts, since drugs represent only 10-20% of health expenses, yet represent a significantly greater portion of the research. CEA needs to be applied to other health issues—for example, hospital treatments, or the everyday things that people do to affect the disease burden (e.g., exercise). Neumann also notes that it is hard to do CEA in the decentralized U.S. health system. It would be easier to do in a nationalized healthcare system. But even with that, there is a fear of rationing.

Neumann believes CEA is preferred over BCA in part because of the nature of the good being valued. In health, doctors and others have a sense that it's immoral or unethical to consider costs and benefits. This is despite concerns about QALYs not being able to account for adaptations of people's valuations over time, and the fact that people tend not to be QALY-maximizers. Regardless, QALYs still seem to be preferred, in part because skepticism of BCA is greater than that of CEA. It also seems acceptable to talk of comparative effectiveness, but less acceptable to talk about cost-effectiveness.

Lou Garrison concluded the panel by asking "Why QALYs and CUA?" He asserted that despite different preferences for risk and different WTPs for health,





medical care is seen as a merit good to which all citizens deserve some minimum level of benefit. Trying to determine WTP for incremental insurance coverage—perhaps the correct test—is impractical. QALYs and CUA are tools to help make consistent decisions about the minimum benefit package. In addition, this only considers cost-access to health, not concerns such as physical access, e.g., through hospital locations and other factors.

Discussion: Barriers to the use of benefit-cost analysis and cost-effectiveness analysis in public health and medical decisionmaking

Michael Adler commented that one difference between CEA and BCA is that BCA relies on heterogeneous WTP values that are income-based. Normatively, to what extent is heterogeneity desirable? Do you want a BCA with distributional weights that may dampen heterogeneity? CEA may actually be a rough approximation of the results you get if you do BCA without heterogeneity. In addition, we need to distinguish between the instrument and its use, and between the instrument in theory and in practice.

Peter Neumann replied that we need to think normatively, and to think about how we conceptualize WTP. The nature of a public good is very different from the nature of a health/medical good, and unpacking and understanding that is useful.

Jack Calfee noted that it's not completely true that a QALY = QALY = QALY. Variation still occurs, though less than with WTP. When CEA is used explicitly, for example, in negotiating contracts in other countries, there is a lot of room for price negotiation and price discrimination.

Richard Zerbe noted that insofar as healthcare and some of its aspects are regarded as merit-based—i.e., things that people have a right to—the correct measure is willingness to accept payment (WTA), not WTP. If WTA is used, the equity differences greatly diminish. If you look at health using WTP values, objections on the grounds of equity do arise; this suggests there is a missing value, namely, what other people are willing to pay for the healthcare of others.

Peter Neumann notes that there is no perfect metric; all metrics will have pros and cons. Measuring altruism is possible, though challenging. The equity challenges against QALYs are based on people's arbitrary views, to some extent. QALYs are not value-neutral.





Lou Garrison agrees that not enough attention is paid to WTP and WTA. Bernie O'Brien has a paper on the kink in the WTA line that is relevant; Mark Pauly's slides regarding Medicare also consider altruism.

Jim Hammitt was struck early in Lou Garretson's talk by the comment that WTP in health is an important argument against BCA. In any public program, there will be some baseline above which wealthier people can buy goods, similar to how education is funded. If we do BCAs with a uniform measure for everyone, WTP may not be a problem. This doesn't address the question of what the minimum should be, however.

Jack Calfee believes there are challenges specific to health, for example, random occurrences of catastrophic accident or illness. When you get to what some people might think is most important—cancer, for example—even a minimum-level health insurance policy is expensive. It's different when relatively smaller costs are at issue, rather than bigger ones.

Peter Neumann notes that it has proven difficult to exclude things from the Medicare package. In addition, controlling costs in healthcare is related to CEA.

Richard Belzer commented that he found the initial premise of BCA being “so much more controversial” in healthcare than in environmental policy amusing, given that in reality, rivers of blood have flowed over using BCA in environmental policy as well. Superfund was a big deal when it was being decided, but now things are pretty quiet, since it's about spending other people's money—not the government's. Connecting this to healthcare, a recent *Wall Street Journal* article talked about how third-party payers were the problem. If we converted healthcare into a public good, the more controversial it may become. BCA is not controversial in areas that aren't covered by insurance, such as Lasik, cosmetic surgery, and veterinary medicine.

Jack Calfee sees a difference between the two—if your dog is dying, the anonymous man on the street won't care. However, if your child is dying, that anonymous man will care. That's the difference.

Reed Johnson believes there are three reasons why the acceptance of BCA in environmental arena is not going to be replicated in the health fields: (1) there is no regulatory mandate to do BCA; (2) physicians act as gatekeepers and are given more respect and have more influence than biologists; and (3) QALYs offer an intuitive substitute for hard thinking, and anything else economists could offer would be more complex.





Peter Neumann notes that the National Institutes of Health and the Center for Disease Control have used QALY studies. Despite their shortcomings, QALYs have been workable. In reality, more than just doctors are gatekeepers.

V. Kerry Smith noted that with scarce aggregate resources, there will have to be rationing of some sort. Whenever there is rationing, there are always criteria that go outside price mechanisms, creating virtual prices. The discussion doesn't seem to be addressing that at all.

Paul Anton notes that “having numbers” is often seen as the high ground, and that having really good numbers is the even higher ground. As economist, he would rather have aggregate numbers and go to convince a state legislator or a single payer versus going to a doctor and saying they should change treatment.

Peter Neumann noted that what we're typically talking about are population-based decisions, public policy decisions, not patient-doctor decisions.

Glenn Blomquist observed that economists have thought a lot about altruism, and that one of the things that's important is what type of altruism you're counting—whether it's general or focused (e.g., parental) altruism. If pure altruism isn't health-focused, you don't add it because it distorts. You would only be able to add health-focused altruism. A study by Per-Olav Johannson suggests there is health-focused altruism; it would be interesting to see if we have that in the US.

Samuel Bonti-Ankomah was quite surprised to hear that CEA is preferred to BCA. Canada has been using BCA as mandate. Cost-effectiveness analysis is recommended as a second option.

Kim Thompson noted that a challenge that has emerged is that interventions are uncertain, variable, and dynamic, but the analyses done of them are not.





Panel II: Barriers to the valuation of health outcomes in benefit-cost analysis and cost-effectiveness analysis

Lisa Robinson introduced this panel focusing on metrics. Many social programs have such profound effects on human health and safety that counts of deaths, illnesses, or injuries averted do not provide a complete picture. There are three common valuation approaches that aim to find a shared scale that allows for aggregation and comparison with other risks and expenditures: Costs of illness (COI); QALYs or disability-adjusted life years (DALYs); and the value of a statistical life (VSL). Each has its benefits and drawbacks. The key issues in this conversation are that preferences can be difficult to measure; constant measures are often used by advocates who assert this promotes equal treatment; and monetary values for fatal risks are widely misunderstood. In addition, we need information on equity as well as efficiency.

James Hammitt presented positive and normative justifications for BCA, including monetary and health-utility measures of health. As a threshold matter, there is the consideration of whether to include expert or public values of inputs. How should the analysis be adopted to incorporate departures from the model? Ultimately, should the analysis be populist or paternalistic?

Fundamental questions for public policy include both individual and distributional queries. On the individual side, what is “good” for an individual? At what level—individual, family, household, community—are we assessing the good? Distributionally, when does more good for one person justify less good for another? Alternative justifications for BCA have included the positivist Kaldor-Hicks compensation test and the normative utilitarian test. The Kaldor-Hicks test raises the question of whether Pareto-efficiency is a sufficient basis for deciding policy. The utilitarian model raises the question of interpersonal comparisons, regardless of whether the comparisons are through money or QALYs or some other measure.

One way to figure out the appropriate threshold for QALYs is to figure out what the social WTP is. For example, does the social WTP depend on the symptoms, type, or name of the disease? Quality-adjusted life years assume the weight of the health state is independent of the duration of that health state, i.e., that the individual is risk-neutral regarding the longevity of any particular state. However, research suggests these assumptions may not be accurate.

Lisa Robinson
Independent Consultant

James Hammitt
Harvard University

Trudy Ann Cameron
University of Oregon

Joseph Lipscomb
Emory University, Rollins
School of Public Health

Dean T. Jamison
University of Washington





REMOVING BARRIERS

In order to explore this, Hammitt undertook a stated-preference survey designed to elicit WTP to reduce the risk that respondent individuals or others would suffer specific illnesses. The researchers hoped to determine whether WTP is proportional to the expected gain in QALYs. Questions were asked about the respondent, the respondent's child, and about other adults in the household.

The results of the survey suggest that there is not a constant cost per QALY. A similar survey, exploring WTP for various short-term morbidity risks found that going to the hospital is roughly twice as bad as being mildly ill but still being able to work. There was a much smaller sensitivity to illness severity. In addition, whether the questions were placed in a daily or monthly framework did not affect responses.

The study concludes that WTP per QALY is not constant across health gains; rather, it is in an increasing concave function of expected gain in QALYs. Quality-adjusted life years are not an adequate measure for determining WTP for health because their value varies based on who in the household was affected, what illnesses were projected, and the length of the illness.

In the positivist justification, this suggests that health measures consistent with individuals' preferences should be chosen, such as WTP or other flexible utility measures. The normative justification of BCA suggests that health measures should be chosen that are consistent with normative principles, such as QALYs.

Trudy Ann Cameron presented on WTP to reduce health risks. Her first point reiterated that WTP is an inverse demand function. In this model, health risk reductions and improved health states are "goods and services." Under utility-based choice theory, you offer people a health risk reduction in exchange for a reduction in income. After choosing a functional form and other considerations, if utility is linear in net income, then WTP depends on the change in risk and the marginal utility of health. Under extreme homogeneity, $VSL = WTP$. But in reality, preferences are heterogeneous, marginal utilities vary, and WTP is heterogeneous.

Studies typically look into small risk reductions across a group of people and measure benefits in physical terms. However, aggregating physical risks is a questionable step when heterogeneity exists. In this way, VSL is an average across estimates from various tradeoffs involving different people. It's just dumb luck if this somehow matches the proposed policy. For nonstandard populations and risks, it's even less likely that you will come up with the correct number.

Cameron's main paper regarding this work posits that the simplest WTP for microrisk reduction depends on an individual's current age, income level, and





discount rate. The model includes interactions that consider “fates worse than death.” The basic model used had unavoidable heterogeneity, and the extensions introduced additional sources of variations.

The study also found that WTP varies by the name of the disease. Even if identical profiles were presented, labeling a profile as certain disease changed the WTP. The WTP for certain diseases was larger than the conventional VSL values for identified diseases, while others were smaller. The study also found that women have a higher marginal utility of income, and don’t want to be less healthy. The presence of small children in respondents’ lives increased WTP, while presence of teenagers lowered WTP. Willingness to pay is also associated with age: The WTP to avoid a lost life-year increases as one ages; however, the older you are now, the lower your WTP currently is to avoid lost life-years.

Cameron’s work also addressed the altruism for others’ health that was raised in Panel I. The survey indicates that when costs are borne through taxes, people tend not to care about others’ health. Willingness to pay still reflects income in these circumstances. This raises questions of the distribution of individual benefits and costs, heterogeneity and paternalism. Should decisions be based on what people want, or what is best for them? Either way, flawed perceptions can’t be corrected unless there is data and a model that permit you to control for individual subjective information.

Joseph Lipscomb spoke on the need for continuous QALY improvement. Many regard QALY as the gold standard for analyses requiring summary measures of health that integrate quantity and quality of life, despite the substantive concerns that have surfaced. These concerns include: Theoretic underpinnings of CEA itself; QALY sensitivity for capturing all benefits that matter; a lack of consensus on multiple conceptual and empirical matters inside QALY as currently determined, such as the valuation of various health states; and how to address equity. A recent key theme is that instead of abandoning the QALY, scholars should work to improve it by learning from other benefit-measuring approaches, such as WTP.

Quality-adjusted life years are the most common yardstick for CEAs, since they permit the combined mortality/quality of life benefits of health interventions to be measured as a composite outcome on a common scale. A QALY is equal to one year in perfect health. In a step function model of QALYs, the order of health states does not matter. In reality, however, health states are multidimensional. The goal is to get cost-per-QALY to be as low as possible, regardless of sequence. Many people have preferences for specific sequences, or for better excellent health





followed by poor health, rather than medium-health throughout.

The debate continues on key methods issues regarding the “conventional QALY.” The conventional QALY is a highly simplified model and does not integrate distributional and other ethical issues. However, to abandon the QALY would sever the link to hundreds of published studies and ongoing investigations tied to significant data. A more productive approach would be to use the standard QALY as the starting point for “continuous QALY improvement.” There are three research areas to focus on in particular: Health state definitions and descriptions; valuation of health states; and distributional and other equity considerations.

Currently, Lipscomb argues, the competing systems of health state definitions and descriptions are difficult to compare since they have different domains and variables. This slows down progress in the field. One solution is working towards a “consensus domain structure” to create a single citizen perspective-based health measurement system. Another option is to cross-walk QALY scores between measurement systems.

Another area ripe for research is valuation of health states. Currently, the derivation of the traditional QALY varies by method of eliciting preferences, by derivation of aggregate QALYs for multidimensional health states, and by integrating “states worse than death.” Cross-walking scores across measurement systems could be useful, or a consensus process leading to a “Reference Case” QALY could be established. Other challenges include the fact that the standard functional form does not allow for health state duration consideration or sequence effects. The accuracy of community-derived health state preferences may also be questioned; perhaps it would be more accurate to elicit the preferences of those who have experienced particular states of health.

Fairness matters, Lipscomb asserts, but it is not a matter that can be settled within the traditional QALY. This shortcoming may be addressed through equity weighting, constrained optimization modeling, or through community-based deliberative processes. In the end, the QALY should not be abandoned, as there is a significant amount of work that has gone into it that we do not want to lose. A concerted effort to improve the QALY is necessary, however, and this presentation aimed to put forward some of those possible areas for improvement.

Dean T. Jamison spoke on financial protection and CEA in health. He introduced his presentation by noting his work focuses on low- and middle-income countries. In the developing country context, the government must sharply limit what can be funded by public sector funds, leading to issues around financial protection that





REMOVING BARRIERS

have shaped the discourse. There are two poles in this debate: Those who are only concerned with QALYs/DALYs, and the community of people who are concerned with health system development and financial protection of the people. In terms of the United States debate, health would much more cheaply achieved by taxing cigarettes, alcohol, trans fats, and the like, but there is the competing concern of financial protection. Low-income countries have focused on catastrophic insurance; the debate of whether to be focused on treatment or prevention has not yet begun.

The domain of CEA needs to be extended. The cost of disease is not just in QALYs or DALYs, but also includes reduced productivity of labor, land, capital, and schooling; welfare loss; and financial risks. In this extension of CEA, what do we learn from WTP discussions? Does this change the view of how to do CEA? What are the economic consequences of better health? How do we aggregate and measure them? What are the costs of intervention? In addition to financial costs, there are health system capacity costs, the costs of medical risks and side effects, the time of patients and their families, and deadweight loss or gain from taxation.

Outcome measures naturally divide into three categories: Outcomes measured in the National Income and Product Accounts (NIPA) (e.g., productivity losses from disease, financial costs of intervention); outcomes that are well-measured but not included in the NIPA (e.g., DALYs, welfare loss through VSL methods); and outcomes outside NIPA for which a range of comparable measures do not yet exist.

In most of the world, the patient bears the risk of payment. What is important to show is the value added when moving from first- to second- or third-party payers. The development literature looks at the consequences of ill health. This is a threshold-based model, either based on income thresholds, or on health-expenditure thresholds. A separate approach is based on utility functions.

Threshold-based indicators include the World Health Report from 2000, and Waters, Anderson, and Mays (2004) “Measuring Financial Protection in Health in the United States,” *Health Policy*. A study with utility-based indicators, McClellan & Skinner’s 2006 paper in the *Journal of Public Economics*, titled “Incidence of Medicare” is particularly noteworthy in that it did not assume that utility was linear with money. If utility is linear, Medicare is slightly regressive. If people are risk-averse, Medicare has good distribution consequences because the value of the risk reduction to the poor has more value than the value of the same risk reduction to higher-income people.

Utility-based indicators for the value of insurance naturally start with theories of risk aversion and intertemporal smoothing of income. The typical utility function





depends on income and the coefficient of relative risk aversion. Financial protection cost-effectiveness (FPCE) is defined in terms of expected value and the certainty equivalent of the gamble based on the probability of getting the disease and the cost of treatment. However, there is considerable disagreement as to the value of this risk aversion, ranging from 0.48 to 3.0. The intertemporality of risk aversion and elasticity are absolutely critical. It is amazing that for this critical value we only have an idea of the value within a range of 5 or 8.

In calculations exploring the application of short-course chemotherapy for tuberculosis we find that the value of insurance falls sharply with income assuming a logarithmic utility function. FPCE is highly income-sensitive. For a person who makes \$200, the FPCE is 36% of the value of publically-provided intervention. For a person who makes \$2,000, it's significantly less, just 2%.

Jamison concluded that in a rough-and-ready way, we should be able to measure FPCE health intervention by health intervention. After this is determined, we could decide whether we want to aggregate. Once these questions are answered, we can start talking about efficient health systems intervention. Given the sensitivity of FPCE to income, the notion of providing financial protection by paying for hospitals is deeply misguided. Low-income people will get less protection through low-cost interventions, but they will also get more benefit. Wealthier people will only get benefits from more expensive interventions, such as hospitalization.

Discussion: Barriers to the valuation of health outcomes in benefit-cost analysis and cost-effectiveness analysis

Richard Zerbe began the conversation with a reminder to consider who is paying the costs. Is your universe a town? The county? A state? When do you rely on public versus individual preferences? Zerbe prefers to rely on informed individual preferences, and if you don't have those, on experts. Is the Kaldor-Hicks test moral? It assumes that the winners compensate the losers; however, the costs of compensation would be so high that it is unreasonable. Zerbe suggests dropping the compensation part of the test entirely. One way to approach distribution weights is to look at WTP for equity (e.g., look to altruism and charitable donations, etc.), instead of deriving it from arbitrary space.

James Hammitt replied he doesn't agree that the Kaldor-Hicks test should be abandoned, and notes that the technical critiques some have raised to Kaldor-Hicks aren't important for most practical situations. If the compensation test is dropped, why add monetary benefits and costs? One





reason is because of the sense that marginal value is equal across people. The better argument is that it is justified by the Kaldor-Hicks test, and that the lower-cost way to redistribute resources is through taxation and welfare, not through explicit compensation policy-by-policy.

Dale Whittington questioned the panel about their optimism regarding BCA's power and use in the healthcare sector, given the long history of debating benefits in that sector. What about in other social policy areas?

Trudy Ann Cameron replied that her hopes hinge a bit on whether or not the discipline is able to get past these labels (e.g., VSL). The shorthand of "monetizing" things is abhorrent to people, even though they realize that tradeoffs have to be made. It needs to be understood that health is not being monetized, rather the opportunities lost when you do healthcare are being monetized. We need to figure out how to talk about things without using statistical lives.

Lisa Robinson noted that the simple fact that someone is out there collecting data and doing analysis has impact on the debate. Just the requirement of having to do BCA can have tremendous impact.

Joseph Lipscomb replied that it seems to be the case that BCA is involved with monetizing the effects. There is literature that points out that if you compute cost-effectiveness ratios, once you make your decision it is effectively as if you have used BCA and drawn your line in the sand. The general tenor is often that CEA is a decision-making tool. Maybe that is all we can expect and all we should hope for.

John Calfee noted that WTP and willingness to be taxed to have a probability reduced are not the same.

Trudy Ann Cameron commented on taxes. The way choice models are set up creates a utility component that derives from lives saved or increased recoveries. There is a lump of autonomous recovery and a lump of negative recovery that goes with each based in part on how the question is asked. In the analysis, we found a huge negative lump of utility. At first we thought it was a programming error. Now I'm worried we are going to find out how selfish we really are.

Glenn Harrison argued that the discussion that has been started with Kaldor-Hicks and hypothetical compensation needs to be ended. In any serious BCA I can always disaggregate and find some losers. When you face the very serious question on how to deal with these losers, you can wave your hands and say the winners could probably compensate the losers. But that can be an invitation to unfunded mandates. In addition, you're assuming





benefits that you may not get, so budget constraints are less likely to be enforced. It also makes programs really easy to shoot down. These concerns are not trivial. There is a large marginal efficiency cost to the redistribution burden. We need to figure out how to redistribute in order to move the discipline forward.

James Hammitt noted that the costs of transfer are a given. If we are going to always compensate losers, we should use Pareto improvement as our criterion. A better way, however, is to adopt a portfolio approach.

Tracy Ann Cameron suggested that we move away from the imaginary notion of compensating everyone with money. Policies affect everyone every day. You need to look at who are the chronic losers, who are the chronic winners.

Cindy Lamy, who described herself as a “consumer” of BCAs, asked about a timeline for moving better measurements forward. Colleagues are always arguing for better measurement, but what today’s discussion implies is that these measurement problems are in fact stopping the discipline from moving forward. In addition, is there a way to specifically apply these mechanisms to poverty populations?

James Hammitt noted that we need to be humble about what we can actually do. There are more uncertainties than just those about value. For example, in the environmental arena, the biggest uncertainty is the dose-response function. That stays the same in education too—what’s the effect of this program, what’s the response?





Panel III: Behavioral approaches to policy

Michael Stegman opened the panel by picking up on Hammitt’s point about real world behavior. If real world behavior deviates from economic theory, is that a problem with the theory or a problem with the people? This panel aims to put that question into a more formal context, exploring implications for policy and BCA of behavioral theories. This is particularly relevant now as there are several prominent behavioralists in senior appointments throughout the Obama Administration, including the Congressional Budget Office, the Department of the Treasury, and the Office of Management and Budget. Cass Sunstein, the director of the Office of Information and Regulatory Affairs, is a well-known behavioralist.

Michael A. Stegman

*The MacArthur
Foundation*

Eldar Shafir

Princeton University

Hunt Alcott

*Massachusetts Institute of
Technology*

Sendhil Mullainathan

Harvard University

Eldar Shafir noted that an overview of the literature suggests the calculus done by people is not very straightforward. There are two conflicting views of the human agent: The rational agent model and the behavioral model. The rational agent model assumes well-informed, stable preferences and a self-interested agent. The behavioral model suggests mediocre judgment, malleable preferences, and the use of emotions in decisionmaking. If the first is true, we can give people a well-functioning market and step away. If the second is true, a more paternal system may be desirable. Benefit-cost analysis, like Pareto optimality, is logically unassailable (i.e., costs and benefits do matter), but the situation is more complicated. When it comes to individual agents, what is a cost? A benefit?

Milgram’s obedience experiments show us that situations matter and their power tends to be underestimated. Decisions are not about objective states, but rather about mental representations of those states. Minor changes in perspective can have a large effect. This has also been shown in other experiments.

Shafir and his colleagues conducted field experiments with advertising for loan applications. The results indicate that not all of the difference could be explained by the cost of thinking. Providing more options to consider decreased the probability of participation. What the default was set at also dramatically changed how people perceived costs and benefits. For example, when explicit consent is required for organ donation, consent rates are low (none above 30%), but when consent is presumed under an opt-out program, consent rates are high (none below 85%). Johnson & Goldstein, *Science*, 2003.

It is not always clear what the “right” policy will be. For example, imposing a fee on parents who pick their children up late from daycare results in lateness going *up*. Now





REMOVING BARRIERS

instead of feeling guilty, parents just pay the fee. The same goes for illegal parking.

Once we accept the empirical view of human nature, it alters how we evaluate policies and products, how we try and help people, and the ethical considerations involved. We need to think about individual decisionmakers not as faulty rational agents, but as fundamentally different creatures whose “construed” costs and benefits are often not what they seem.

Sendhil Mullainathan explored what Eldar Shafir’s talk implies for policy design, and for BCA of both traditional and novel policies. He described three styles of policy design where this has an effect. The first, nudges, explores the idea that defaults have a lot of cache in policy circles. Similar to the organ donation default Shafir mentioned, we see differences in 401(k) enrollment and default for employees who joined a firm prior to automatic enrollment versus those who joined after. Here, the transaction costs of opting in or out are tiny, yet the differences from default are astonishing, not just immediately upon joining firm, but long-term and going forward for years. Nudges can have enormous power in moving effective, long-term behavior.

A related type of nudge is how you provide information in how you frame the disclosure. For example, how should the mandated disclosure of fuel efficiency be phrased? Miles per gallon? Over 10,000 miles, an improvement from 10 to 11 mpg results in the same gas savings as moving from 33 to 50 mpg. A similar framing problem comes up with buying a calculator versus buying a computer. People were willing to travel the same distance for much less of a monetary discount to buy a calculator than a computer. Even though the value of a person’s time really shouldn’t change, the base price determines the “optics” of what your time is worth. What framing effects should we consider when designing a disclosure policy?

A final example is choice architecture, which is a little more on the frontier, but is gaining traction. We tend to think choice is a good thing: The more choices, the more you can sort to match your preferences, and when people are choosing well, firms have incentive to compete for people’s tastes. But if people are choosing poorly, both of these fall apart: You won’t maximize your preferences, and firms will just maximize their profits. Consider Medicare Part D, one key feature of the program is choice. We looked at whether seniors made the right choice given their basket of drug utilization. The answer was no: The average cost of the plan selected was much closer to the median price than to the least expensive plan.

But how much of this difference was about true preferences, and how much was a mistake? If we assume people are rational actors and use stated preference models, we could just say that the difference in costs





REMOVING BARRIERS

measures the value of the plan to those seniors (ease of using the plan chosen, etc.). For half the sample, data on program differences was given back to the senior. Even though this data was publically available on the website, giving that data back with no real framing increased seniors' switching plans by 12 percentage points. If the policy were to automatically send seniors their plan data year by year, instead of letting them to go to the website to compare, what advantages might we find? In short, how ought we structure choice to make choice as easy as possible?

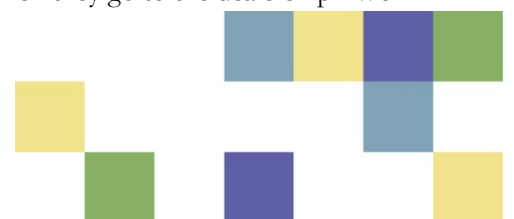
Shifting to discussion of BCA, we run into the challenge of inferring a hedonic value. For example, we assume people buy potato chips according to a benefit-cost formula. But what if the formula is different because people are overvaluing benefits or undervaluing costs? How do we measure the costs and benefits of a tax if we do not know people's formulas for costs and benefits?

Similar challenges raise moral hazard concerns with health insurance. In a traditional model, a co-pay induces people with small benefits not to go to the doctor. In a behavior model, people sorted out through use of a co-pay may actually be those who would get the highest benefits from going to the doctor, only they are lazy, or procrastinate, or some other reason. Elasticity of demand is not a sufficient statistic. In short, we need new tools to incorporate these ideas systematically into BCA. We need new field experiments and to measure more and different things.

Hunt Allcott applied some of the same concepts to energy and climate change policy. In particular, he focused on the standard BCA with rational consumers, evidence of consumer inattention to future energy costs and its implication for BCA, and how to use field experiments to advance BCA in this area. For example, when consumers think about an SUV, they include the costs of the SUV and gas, but they tend not to think about the externality of climate.

Hausman (1979) noted that people have implicit discount rates that are much higher than the cost of capital: We have a defective telescopic faculty. That is, it is hard for us to think about future costs.

Looking at used vehicle markets, you might think that demand for Hummers versus Priuses should have changed, showing up in market share and resale value. The price difference between fuel efficient cars actually does track gas prices, but there is a substantial mispricing in the used car market. Imagine a \$1 permanent change in gas prices: Actual changes in relative prices for a 40mpg vehicle versus a 20mpg vehicle are only $\frac{1}{4}$ of predicted changes in relative prices. Why? Because people don't think a lot about the history of gas prices when they go to the dealership. We need more empirical evidence to prove this, but





assuming it is correct, what does it mean for BCA?

Under a carbon tax or energy efficiency standard, we can calculate a gain from consumers who bought SUVs but should not have because they mis-optimized. We can also calculate a consumer surplus loss from those who could buy optimally and now cannot because of the standard. But how do we know where to set the fuel efficiency standard? Do we assume that everyone has the same level of inattention to gas prices? In fact, people pay different levels of attention, so a bright-line fuel efficiency rule is not going to maximize benefits. In the end, choices do not tell us anything about welfare.

Discussion: Behavioral approaches to policy

Glenn Harrison noted that in the context of medical decisionmaking, there is the argument that you should not apply the maxims of prospect theory, and should not confuse descriptive models with prescriptive policies. How does that fit in with these discussions?

Sendhil Mullainathan replied that there are two separate issues: (1) what policy levels are available if we frame consumer surplus loss as a gain or vice-versa; and (2) what hedonic value do we assign when the loss works better than the gain. The challenge is to find the context in which we can overcome the problem of hedonic measurement. The deeper issue is that if we see evidence that poses problems for the model, we cannot ignore it. We have to be convinced that this is how people behave, and we have to move from the lab to the field to see if this is true.

V. Kerry Smith noted that economists do not know preferences, but we do know choices. If we keep track of the way that material is presented to people, we can always adjust and reconstruct as we learn more. For Medicare Part D, did things change as the materials improved?

Sendhil Mullainathan replied that the original data is from 2007. The data from 2008 and the follow-up do not suggest that there was much improvement. However, it is possible that in the long run it will improve, though other studies (e.g., on mutual funds) show persistent confusion. The market has some incentive to become simple, but the market also has other incentives.

Richard Williams stated that most of the research he saw applied to how people deviate from rationality. He would like to see the same tools extended to how people make decisions for others, e.g., irrationality in government decisionmaking.





Unleashing the Power of Social
Benefit-Cost Analysis:

REMOVING BARRIERS

Eldar Shafir noted that there is a general tendency to assume that decisionmakers are fallible as people. But in most subjects, when public decisionmakers are shown what they are doing is irrational, they adjust their behavior. Where important aspects are made salient, or where people can consult with experts, they might do better. It is not necessarily true that irrationality is unavoidable at any level.





Panel IV: The use of benefit-cost information in education policy—barriers and opportunities

V. Kerry Smith
Arizona State University

Clive Belfield
Queens College

Tom Lamb
*PNC Financial Services
Group, Inc.*

Michael Rebell
Columbia University

Al Lindseth
*Sutherland Asbill &
Brennan, LLP*

Convened by **W. Steven
Barnett**, *Rutgers
University.*

V. Kerry Smith noted that Steven Barnett convened the session, and presented briefly on work he is doing with Klaiber on adapting BCA for general equilibrium effects of social policies. The project explores school quality and housing price effects in Maricopa County, Arizona. The central focus is the non-market feedback effect of housing prices.

Housing prices are related to K-12 school quality. If school quality exogenously changes, then it is possible to set up an equilibrium model, albeit a different kind than economists are used to. In Maricopa County they fired teachers for budget reasons but the students are still there. Class size is endogenous, families moved, and housing prices changed. Once the equilibrium model is set up, it is possible to estimate and perhaps solve. The model they are using is a sorting model. Maricopa County has 46 school districts. Controlling for school characteristics for every district, they looked to the contribution school distance made to school prices, and were able to determine a set of fixed effects from about 15-18. The price by district is equal to the contribution that school district makes to house prices after controlling for everything else.

Klaiber and Smith were able to rank school districts, and those rankings matched almost exactly with standardized test scores. The prediction from the model is that school district quality would match with housing prices if housing prices matter to school district quality. Based on the parameters used, and the data on test scores and student-teacher ratios available, they found that if more students come into a school but the number of teachers stays the same, school quality as measured by test scores will go down. If you fire teachers, the quality of education goes down, families (and children) move and will keep moving until equilibrium is found.

Clive Belfield noted that practice is a long way from theory in education. There is a need for better costs analysis in education. Part of the challenge is that the burden of funding is both public and private, comes from multiple sources (states, local agencies, federal grants, parents), and involves weak market forces. In addition, the costs are paid long before seeing the benefits.

The current practice of cost analysis in education is essentially to downplay it. This comes in several forms: Focusing on curriculum reform, which “doesn’t cost anything;” working to improve teacher quality





REMOVING BARRIERS

because “we just have to;” or downplaying costs overall. The United States voucher model asserts the analysis is unnecessary because either the programs are so wonderful or it’s apparent that private schools are cheaper. Other studies use “back-of-the-envelope” calculations when a lot of money has been spent on impact, trying to recalculate costs using a simplistic framework long after the program is implement. Yet another approach is to assume that the private cost of education is zero. In total, we are not doing a good job of measuring costs.

Belfield asserts this situation exists because cost analysis is too easy and trivial and yet too inaccurate or expensive. The one exception is perhaps early education, but the same programs keep coming up, and these programs tend to be those that work with a very restricted group of people. In addition, oftentimes schools of education that do not have economic knowledge are doing the analyses. We are far from agreed-on standards for BCA in education. Most analyses infer their cost analyses, and do not include opportunity costs. They also do not include return on investment or net present value.

The focus on benefits is not much better. Often studies focus on “impacts” rather than benefits. Is one explanation just a sense that “costs analysis must be difficult?” What about the lowered incentives for researchers who sense that costs analyses and BCA are not valued?

Belfield proposes a “workable accord” on methods to estimate cost per unit output of a given quality. This accord would determine the “ingredients” of an analysis and the accounting method, and utilize a cost-function method. Potential problems with this accord include unconstrained, complete input use. Other challenges include the “hidden inputs” of education that have unknown prices: in-kind resources, grants or transfers from external sources, program design and participation or compliance, and students’ time and effort, to name a few.

On the production side, it’s a lot harder to know what inputs are needed to produce an output unit of given quality, particularly given that multiple inputs can create the same output. However, for BCA it does not really matter *why* things work, the analysis can still be applied. It is also hard to specify outputs: Should we be looking at academic or developmental outputs, for example, or maybe at adult earnings?

On the whole, BCA in education is not valued because of practical objections, uncertainties about cost data, no cooperation, limited funding to do analyses, and the sense that it constrains policy decisions and provides “even more opportunities” for policies to fail. In addition, there are few BCA courses taught in graduate schools of public policy and education—many do not know how to use this tool.





REMOVING BARRIERS

In the end, how do we unleash BCA? A few options include doing *ex post* analysis on existing BCAs, or coming to agreement on shadow prices for present-value benefits.

Michael Rebell shared his experience as an attorney looking at what happens when very important legal and policy decisions around education are made in situations where there *is not* sophisticated BCA data. In particular, he noted the recent wave of litigation challenging the constitutional sufficiency of education financing in virtually all states. The United States Supreme Court has held that there is no constitutionally required level of funding for education in the federal constitution. In the past 35 years, there have been cases in 45 out of 50 states raising various issues related to education funding.

Campaign for Fiscal Equity vs. New York challenged the fairness and constitutionality of the education funding system in New York state. One of the most fascinating things to emerge from this litigation was that when you put a microscope on how the state legislature had determined how much money to allocate to districts in the state, the determination was that there was no analysis. Specifically, the court was shocked to realize that \$15 billion was being allocated by the state and no one had undertaken any kind of cost analysis to determine what the needs were. No one had considered whether more money should be spent on low-performing students, gifted and talented students, or students with disabilities. Trial was able to prove how the formulas operated. Trial also showed that a political deal had been made 10 years before that New York City schools—which claimed they were underfunded—would receive a slight increase in proportionate funding each year. So, for example, if they initially were getting 36% of the allocation, they would get 38.86% of the *increase* over the next given number of years. This increase was consistent, despite the number of moving parts in the funding formula.

This was a clear back-room political deal. In order to make the formula work, schoolchildren were being counted as 0.86 of a student. When this came out in court the court declared it unconstitutional. The remedy the court established was that the state had to undertake an analysis of what the cost actually is of providing a sound basic education. Once that was determined and defined, it was logical that the price would be figured out. Despite Belfield's feeling that no one put any value on education programs, the court essentially ordered the state to put a price on it.

The plaintiffs in these school funding cases have won about 60% of their challenges in cases across 45 states. In the process of courts looking at education funding, the education industry has developed new approaches to cost analysis. Though it has elements of BCA, it is not formally BCA. As a





REMOVING BARRIERS

result of these court cases, more than 35 states have undertaken various types of cost analysis for determining disbursement. There are four broad categories of methodology:

- *Professional judgment.* Panels of teachers, business officers, and administrators decide the inputs that are needed to meet the output, usually using a “rough judgment” approach.
- *Expert judgment.* Based on literature of practices that seem to have worked.
- *Cost-function approach.* Based on econometric and relationship analysis, looks to the relationship between inputs and outputs.
- *Successful schools.* Looks to districts that have been successful and asks what they are spending. Assumes that any school that received that amount of money would be able to be successful, given adjustments for poverty, students with disabilities, etc.

Each of these models has varying strengths and includes elements of BCA. When courts started looking at these questions, certain defendants (often states) objected that the court should not be ordering any more funding for low-funded districts because cost study analyses showed that the amount of funding didn’t matter. Through meta-analysis, the states were able to show that by and large, pouring money into school districts didn’t yield an increase in results.

This in turn led to a counter-literature arguing that money did matter. It is a common sense expectation that if you lower class size or increase teacher quality, you are probably going to get better results for students. Most of these court cases have battled out the question “does money matter?”

Judges in about 30 states determined that money does matter. Their decisions were not necessarily based on economics, rather, they cited a political angle, asking why rich districts were trying to protect their allocations if money didn’t matter. In reality, of course, money does matter, but it has to be spent well. This is especially important when we do not have the luxury to spend. We need to look more closely about how we are spending money to be more efficient and effective.

Al Lindseth spoke on making education dollars more cost-effective through performance-based funding. In the area of education funding, there has been little or no BCA applied—and this includes the New York legislature. The state legislature has done well in providing money for education. Over the past decades, education spending has increased fourfold, but we’re not seeing the effects. There has been little or no increase in reading and math abilities for 17 year olds, and an actual decrease in science achievement levels.





REMOVING BARRIERS

We need to spend more money on education, but we need to do it effectively.

Lindseth doesn't believe that costing-out studies are the best way to make effective use of limited education dollars. As an example, he describes Wyoming and the National Assessment of Education Progress (NAEP) trends. Wyoming is one of few states that had enough money to fund studies and recommendations. It went from being a state that spent a middle-of-the-road amount to spending among the highest amount. But if you compare how Wyoming moved relative to the national average in achievement tests, it actually regressed in some areas. Wyoming also was not doing well compared to neighbor states, despite spending more than them on education.

Why haven't we gotten more effect out of our education spending? One reason is that decisions about policy are made separate from funding decisions. We need to have a policy—a way to fund education—that makes funding work for us. Rather than simply paying for programs, we need a performance pay system that encourages achievement. We need a system where the state adopts standards, but local decisionmakers are empowered. Most of the time, policies are dictated by the state. Some district representatives say that more money does not matter; they just need fewer restrictions. Policies need to institute incentives and rewards for success, including merit pay for teachers. The consensus in the research is that academic degrees and the number of years teaching—the standard bases for determining teacher pay—do not have any effect on improving student achievement, except perhaps in the first few years of teaching.

If we are going to take strings off of local districts, there need to be alternate ways to hold them accountable for results. We need to incorporate evaluative components into programs and use rational and equitable systems to fund.

So why aren't costing-out studies the solution, Lindseth asks. He recognizes that they have some use in that they are informative to the legislature to take into account as one factor among many, however, he doesn't believe they should be used by courts. For one, notions of cost-effectiveness play a minor role in these studies. Professional judgment panels are often specifically told not to consider how much the program is going to cost and where the money will come from. This results in essentially a wish list, rather than something that looks at the most cost-effective ways of delivering services. On successful school studies, there is little effort to distinguish between schools that attained success with lower costs, versus those with high costs.





REMOVING BARRIERS

These models usually do not consider new types of education or structures such as charter schools or merit pay. For panel-based models, who appoints the panel has an effect on the biases of the panel. There is no logical reason why costing-out results in 18 states should be so uneven, yet every time you try, you get a different result. Costing-out has not been proven to be reliable in helping to guide education funding decisions.

Tom Lamb noted that he found it interesting that this panel followed a panel on behavioral economics, particularly as we look at how to do behavioral economics with respect to rationality. Even with a quest for rationality, things have become less and less rational. Everyone evaluates through his or her own prism. Lamb noted that he wasn't here to present on how to construct a BCA. He's here to talk about how to use these analyses once they've been created.

Lamb works for PNC Financial Services Group and Bank, the 5th largest commercial bank in the United States with branches in more than a dozen states. About 5 years ago, PNC committed to spending over \$100 million in early childhood education over the PNC footprint. A big part of that pledge is a commitment to advocate for support for increased public and private funding for early childhood education.

PNC takes this commitment very seriously, and has engaged in deliberations in state capitols with respect to funding of early childhood education. A variety of the studies that have been mentioned here have been very helpful in advancing the cause. PNC has had some success in getting funding for early childhood education where it wasn't funded previously, and in holding gains during budget cuts. This success was in part because of these studies. So when do these studies seem to be successful and persuasive?

The first barrier is fundamental politics. At the end of the day, the requirement and challenge is to get a certain number of votes, and you have to do whatever you can to get that number. If BCA helps, great; if it doesn't, then we'll get rid of it in that situation. In education, a lot boils down to incrementalism. The funding was set years ago, and the only way to get support is to bump everyone up a bit, and not to bump anyone down. Another challenge is whether or not the analysis aligns with a particular decisionmaker's values, whether it matches her existing theory or preconception.

The second barrier is comprehension. For lay people, BCA often is not understandable. Public officials do not really know what they are looking at and do





not understand the analysis. The third barrier is whether or not the analysis answers the question that is being asked.

In any kind of radical change in public policy and investment, there is the issue of the time horizon—short-term versus long-term makes a big difference. Benefit-cost analyses typically talk about value earned over a period of time, usually quite a long period. In the business setting, that is important, but in the political realm, it is not as important, because legislators are often thinking in months, not years.

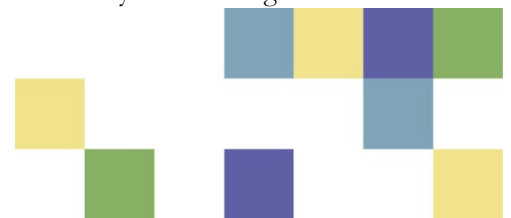
In this process, PNC has learned several lessons. The first one is about messaging. One of the great developments of the decade has been the growth of economic analyses. The concept of return on investment (ROI) is important. People have been arguing forever about future benefits, but ROI numbers are understandable to legislators and business people.

Longitudinal studies are also useful because they are easier to translate. This is setting aside issues about how to calculate the numerical outcomes of those studies and whether the ROI is comparable between studies. Longitudinal studies provide a demonstrated value, and that is easier for decisionmakers to understand. This illustrates an ongoing challenge between the need for academic and scientific rigor, and the need to keep things simple. To be useful, an analysis must be credible and understandable. Outside endorsements will increase the value of a study.

There is also some question about willingness to use BCA, even though perhaps in this economic climate it is more necessary than ever. In a situation where every spending decision is a difficult one, people are trying to discern which decisions make the most sense. It is important to be talking to public officials on the front end about what questions are being asked, and to understand the environment in which they are operating. Another wrinkle is that the real public policy discussion going on is not always the one that is being reported in the media. Regardless, business saying “we understand this, and we stand behind it” has great leverage with the legislature.

Discussion: The use of benefit-cost information in education policy

Kim Thompson noted there has been very little structured analysis about how the education system has changed as a whole, for example, the growth in special education. We need to make sure cost analyses are done consistently while also looking at trends over time. For example, analyses now may be including students





who were previously excluded. Technology has gotten into schools, and that is adding to costs. It is good that we are starting this discussion, but it is challenging for a lot of reasons.

Michael Rebell noted he disagrees with Lindseth's statement that costs have gone up. Demographics have changed. Even if performance was the same as it was in 1960, school systems have achieved a lot to keep the same level of performance given the significantly changed population.

Al Lindseth replied that the demographic changes had been looked at. Special education accounts for approximately 20% of additional costs.

Glenn Harrison noted that in the context of education, there are a couple of 600 pound gorillas in the room including teachers' unions and white flight.

Richard Belzer noted that it is risky to use aggregates and compare them. Aggregates disguise all that is interesting about what is going on. He also noted that the federal regulatory arena has taken a look at the quality of information actually used. There was an initiative in 2001, finalized in 2002, that required federal agencies to take a serious look at the quality of information before disseminating it. Agencies do not like to do this, in part because when they start looking, it sometimes turns out the data quality is quite bad. As economists, we are trained to take data at face value and assume that it is correct. As information gets challenged, economists may be caught in a position between attorneys and legislators.





Panel V: Estimating the willingness to pay of the American public for childhood poverty reduction

Dale Whittington noted that this panel was different from the others. It is focused on work for the MacArthur Foundation that is just starting. Whittington and Johnson presented their collaborative work, and the panelists commented before opening it up to general comment.

The central question of the project is “What would it be worth to you to live in a United States without poverty?” The project involves respondents from all over the country, and starts with a choice between two hypothetical countries. The countries are described in terms of environmental quality, poverty, healthcare, and annual taxes. Measures of poverty include the percentage of children in poverty, high school graduation rate of children in poverty, and pregnancy rate of high school teens. Respondents have not yet been told that the study is about poverty; the design is meant to allow other things to compete with poverty.

The study also involved a contingent valuation scenario where examples of programs were given and respondents were asked how they would vote on legislation before Congress if it would raise their taxes. Results suggest that people care about how the programs were delivered. Within social priorities, poverty was a mid-level concern, though many issues that are related to poverty (e.g., healthcare) also scored high.

When respondents were asked how to reduce poverty, many people said through education. One of the most surprising things was how close poverty was to many people, in terms of themselves or relatives having been in poverty. Many were afraid of falling back into poverty—we saw this in middle-class people as well.

In the legislative choice questions, we had a large number of people who voted “no” to supporting poverty-reduction legislation. Over 40% of respondents said they would not support it at a \$100 tax level. In contrast, 14% said they would support it even at \$10,000. When we looked at these results by political ideology—determined by a separate series of question—a few things really jumped out. Even at \$100, only about 30% of conservatives would support the policy. Respondents also cared a lot about “breaking the cycle of poverty.” That label alone means a lot to people, and they considered education a way to do this. The main reasons people voted against the legislation were because they thought that other government programs should be reduced, or they felt they couldn’t afford the

Dale Whittington

*University of North
Carolina at Chapel Hill,
Manchester Business
School*

F. Reed Johnson

Research Triangle Institute

V. Kerry Smith

Arizona State University

Vic Adamowicz

University of Alberta

Glenn Harrison

*J. Mack Robinson College
of Business, Georgia State
University*





taxes. When compared to the estimates of WTP to reduce poverty, the WTP is less in the legislative scenario than in the country choice scenarios.

F. Reed Johnson noted that the numbers are implausibly large. If we multiplied the WTP by the number of people in the country, we could solve all our problems. So why the yea-saying? There are strategies to reduce yea-saying. We asked people whether they thought poverty would be substantially reduced 50 or 100 years from now. Most people could not conceive of a country without poverty. Many thought poverty might get worse, but they did not see how to fix it. It was something they cared about, but did not know how to fix.

We questioned people on whether they would donate time or money to solve poverty. In focus groups, we were struck by the amount of volunteering and community involvement. A lot of people have experience and ideas about what works. Perhaps this explains some resistance for taxes. At this preliminary stage, the takeaways we have are that people place a high value on reducing child poverty and care about the means to do so. Conservative respondents also placed a higher value on encouraging time donation and local initiatives.

V. Kerry Smith noted that Whittington and Johnson do fascinating work and take on big problems in very creative ways. The first thing Smith wants to start with is defining childhood poverty. It can mean many different things, but the structure they are using is that eliminating poverty is a public good. One thought might be to think of childhood poverty as a condition of a set of children today. However, that tends to get people to think about quick-fix, transfer type programs.

An alternate way to conceptualize childhood poverty is by talking about process. Do people think it will be here 50 years from now? The object then becomes how to transform social economic structure in a way that gives rise to a positive change in the fraction of children in poverty. That is a different public good than that associated with transfer.

Another concern is the object of choice, and that that affects respondents' choices. We need to be very clear that the ultimate goal is to develop a measure of aggregate tradeoff that is associated with allocating resources to eliminate child poverty. They are talking micro-strategy, establishing tradeoffs at the household level to estimate the market for eliminating child poverty, and then aggregating them. An alternate approach could be taken, similar to Heckman, categorizing the tradeoff as avoided "bads," and increased "goods." This uses an effects orientation, with the same goal.

Whittington and Johnson's analysis strategy involves both constructed and elicited





REMOVING BARRIERS

tradeoffs. When we think about listing tradeoffs, we need to think about actual and perceived baseline expectations. People's experience with the object of choice will affect the overall plan, and each element must be evaluated. For example, if you think of the input as a transfer program, then the tax payment vehicle takes on a different role, because what you are looking at then is the permanent transfer of income from one set of people to another set of people. Respondents may look at that very differently than they look at a commodity. Likewise, setting the baseline conditions is important.

With regard to respondents preferring to donate time rather than money, there is the issue of commitment. Oftentimes, people say they're volunteering, but they're not actually doing it.

Smith looked at data from the Tax Foundation to make a quick comparison to figure out what people are currently paying net in federal, state, and local taxes. According to these estimates, low-income groups would be getting \$35K in benefits. In contrast, the higher income groups are paying \$48K on a net average. If you look at gross taxes, the number is hugely different. Many people are getting transfers, and others are paying. How respondents approach the questions in the survey will be different depending on which group respondents are in.

Using a simplified algebraic model, we can illustrate the effects of the plan itself on WTP to pay for the plan. It is important to note that this is a *plan* and an allocation of resources to reduce poverty—it is not a guarantee. It will not dollar-for-dollar reduce poverty. If we calibrate one of the parameters to match Whittington and Johnson's plan, we can calculate the measures of elasticity of substitution. If we start with 2,000 units of poverty and reduce that by a certain amount each time, we see that that if the substitution elasticity goes down, the program gets *much* less effective. Respondents are likely to be far more sensitive to the elasticity of substitution; therefore, how you describe the plan is critical. Smith would like to see a structural model developed with the stated preference survey, and have benchmarks to see if people appreciate what is being asked.

Smith then brought in Heckman's structural model of child poverty, which suggests that a person's earnings function is a product of their cognitive development and education achieved. These wages then go to the household, which shares the income. If we are interested in the cognitive development of children, we have a certain income-sharing rule. The virtual price placed on reduction of child poverty is a function of what they themselves realized through the effect of earnings.





Vic Adamowicz noted how difficult it is for people to conceptualize poverty reduction. There is a lot of baggage that goes along with these questions and that makes it hard to understand what the responses are. Is this something that people would actually vote for if faced with a referendum? Does altruism count?

Adamowicz thinks that the pilot already has a lot of high-quality investment. Trying to triangulate through multiple methods is strong. The notion that people are rejecting the scenario means that they are perceiving something else—it will be important to figure out what that is, and to what extent it affects the model. A mechanism to address the hypothetical bias is necessary. This needs to be dissonance-minimizing and allow people to make socially uncomfortable statements.

Adamowicz’ concerns and suggestions about the pilot fell into several categories. First was the question of whether respondents thought that a 50% reduction in poverty was possible in the time horizon given. A scope test on that would be useful. In addition, the geographic distribution of poverty in the United States is not uniform. We may expect people to respond differently based on poverty in their area. Again, it would be helpful to have a scope test to explore this.

There is a lot of concern about “cheap talk.” How do we remind people that we want them to focus on real money? Perhaps a dissonance-minimizing question design that allowed people to make “maybe” choices would be more realistic. It is difficult to determine whether respondents are rejecting the scenario design, or the social desirability of poverty reduction. Eliciting a perception about how likely the program is to be accomplished may help with this problem.

Adamowicz noted longstanding concern over stated preferences. What are preferences? How can we count them? Levitt and List break up individuals’ utility into moral and wealth components. Is it desirable or possible to isolate moral and wealth components? Does this remove the social desirability effect that would not occur in an actual voting or choice situation? As a separate concern, how well can individuals predict their own behavior? Recent scholarship suggests that individuals in some contexts are not very good at predicting their own behavior. Sometimes asking about what others would or should do gives a better idea of what people themselves would do.

Another consideration is whether there is a preference for the status quo. Trudy Ann Cameron’s paper mentioned a “lump of utility.” Usually this lump of utility is positive, however, people tend to stay at the status quo when things become more complex. One approach is to build the structure and then ask the questions on why





respondents are staying at the status quo, if that is what they are doing.

Adamowicz suggests that the next steps for this project might include modifying some of the survey questions. In particular, the country choice questions may be difficult for people to conceptualize and place themselves in. It may also be fruitful to consider decomposition of the status quo. Adding time to think may be an important element—people need time to ponder the choice they are about to make.

Glenn Harrison looked at challenges with using contingent valuation surveys, with evaluating the risk of poverty and “vulnerability,” and normative challenges. Contingent valuation surveys often have high hypothetical bias. Dichotomous choice, referenda-style, and conjoint choice models all still show bias.

A few solutions using lab and field experiments include ex ante or ex post calibration. For example, if you know that men always exaggerate but women always tell the truth, you can adjust for that after the fact. The challenge then is to find the calibration value.

Other potential issues with the survey include the fact that it is a web-based survey, the lack of a “no choice” option, and the budget estimates. Web-based surveys typically have small sample sizes and may have significant selection bias. Including a “no choice” option changes responses dramatically. In the current survey, the estimates and budget constraint may not be plausible. For example, a WTP of \$3,800 is 6% of total income for a person who makes \$66,000 a year. For an average person making \$66,000 a year, \$3,800 is equivalent to 50% of the food at home budget, 100% of the eating out budget, 400% of the alcohol budget, 150% of clothing budget, or 100% of the entertainment budget. Is this realistic?

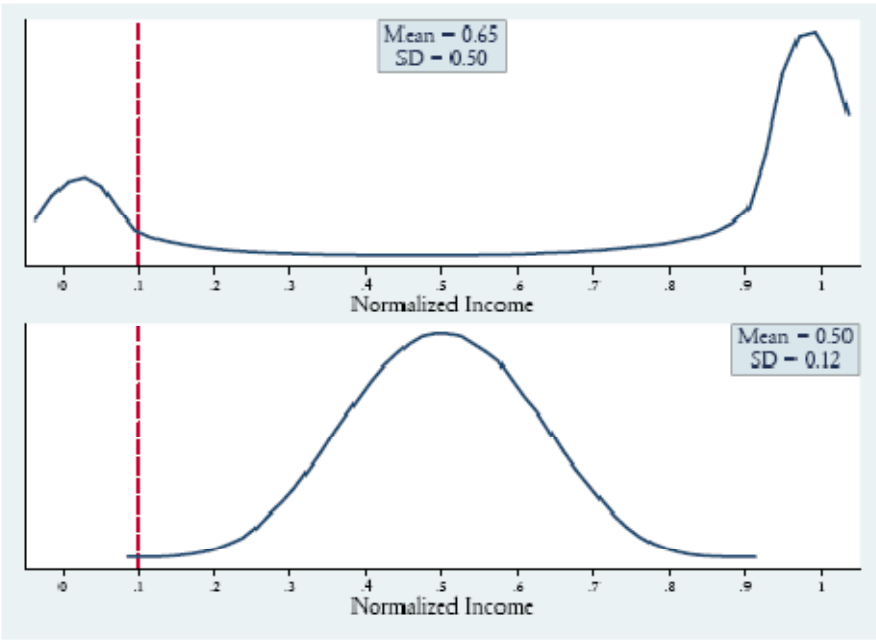
Harrison feels that surveys are “dinosaurs,” and does not feel that scope tests always increase their validity. With both, there is the danger that the “Randomistas” will come after you: Some feel you need field experiments in order to have “real” data to bring to policy makers. Surveys should complement other tools.

Harrison also pointed out potential challenges with thinking of “poverty lines.” There is a lot of fear of getting into poverty, and there is a different fear around short, episodic poverty, versus long-term poverty. We need to consider the whole distribution. How people will answer will depend on how risk averse they are. For example, in the diagram below, the average income is higher in the first graph, but if you are risk averse, you may prefer the plan that reduces average income if it also reduces the probability that you will be in poverty. In short, it is essential to look at the whole distribution.



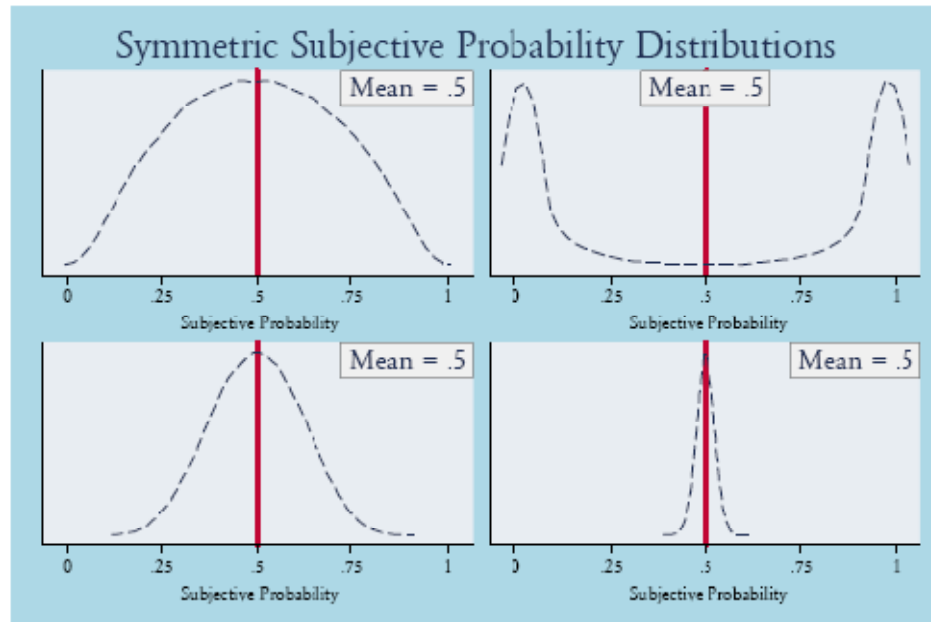


REMOVING BARRIERS



In addition, when we evaluate poverty from the eyes of those who are in it, we need to know their risk attitude; their aversion to uncertainty. If you are risk-averse and your beliefs are based on the reduction of compound lotteries, you value the four graphs below the same.





But it cannot possibly be true that these distributions would be valued the same. They are not the same, because we have not yet taken into account uncertainty. Being risk-averse and uncertainty-averse are different. In terms of policy, if your gut tells you that you would care about which of the above lotteries you would prefer, then you have an aversion to uncertainty. Understanding uncertainty-aversion is critical to understanding risk-aversion and poverty.

Another area in need of analysis is household utility. Are households cooperative? Non-cooperative? What do our beliefs about the altruism of the other person indicate? What is the net analysis if I give money to my daughter, but she cares about my well-being and so is troubled by my having less money?

Discussion: Estimating the willingness to pay of the American public for childhood poverty reduction

Kim Thompson noted that it was important to interface with the community of people who are working in maternal and child health. In part, this will help understand what the feedback loops actually are. The household issue is sticky. You have to start thinking about family planning and questions that are politically and socially very challenging. Focusing on Heckman's theory may be too simplistic; we may also need to look to Holzman's theories. The real challenge is to do more work, first on the system that exists now and its





incentives, and then on what the interventions are through BCA and what questions and surveys we need to do to get the information we need. The maternal and child health community will be interested because they are looking at why people aren't willing to pay to reduce poverty or child poverty.

Glenn Blomquist reiterated the skepticism toward volunteer time, even noting that there *is* a lot of volunteer effort. Maybe looking at time-use or budget studies would access different information. He also noted that reducing poverty by 50% may be a proposition that strains credibility, therefore undermining the questions.

F. Reed Johnson replied that they have struggled to figure out how to frame the questions. There seems to be skepticism about the government's mechanisms, and more faith that local efforts are going to be more useful. People do seem to have strong thoughts about the process of poverty reduction.

Sara Watson commented that she would be interested in how the results of the study would be used in a policy or advocacy environment.

Dale Whittington would like to make the case in BCA terms that reducing poverty is something people care about, and that in relation to its costs, it makes sense.

F. Reed Johnson noted the reason he and Whittington are doing this project is to determine whether people are really willing to pay for poverty reduction up front. Either they do not really believe that poverty reduction is possible, or they care about the process. They are not only trying to get outcomes.

Glenn Harrison noted that Heckman's equation shows there are net benefits. If we have a classic social welfare function, that is great. But "do we want to do this" is an important question. Even if it is politically unpopular to say, we may just not care about poor children.

Kim Thompson noted that the framing effect is significant. What are people thinking about when you say "poverty," versus "poverty reduction?" How does this compare to what they are thinking about healthcare and environmental concerns. We may not care about poor people, but we may care about children. For example, Maureen Cropper's work is about helping children. But people in the maternal and child health field say we cannot reduce childhood poverty without looking at the household. Right now we are not sure how people are responding to the question. Do they see it as a one-time cost to eliminate poverty forever, or a price to be paid every year to continue to eliminate poverty. This will change some people's answers.





Mark Long noted that when you talk to MPA students, many of them believe poverty is more relative than absolute. They believe there will always be poverty, because we will redefine what poverty is. It is a framing issue. Do people believe there could be a 50% reduction in poverty? Should there be? Should we always identify some people as poor?

Michael Stegman spoke to MacArthur's interest in this project. Early on, internally they stopped using the "reduction by 50%" language. One reason why we wanted to do this study is the methodology. The larger goals of this program are to bring significant social policy into the methodology and have people really struggling with the incredibly difficult issues in domestic social policy. From that perspective, this project cannot fail. In terms of Sarah Watson's question, we view Heckman's work as an early education strategy. We were trying to think about how you could use this in another area where there were significant building blocks of evidence that individual programs could show significant net benefits, but all of the BCAs are about small interventions. Most of them are not to scale, and have very high net benefits and benefit-cost ratios. How can we make a contribution in this area? Can we apply this methodology to a broader policy within which one could show that we have a set of effective tools to address issues with various evidence-based programs? The question then becomes are you willing to pay to support such a strategy? If all BCAs are built on costs avoided, but people do not receive or value those gains or costs avoided, then programs can never get to scale. For example, Kerry [Smith], your last chart showed that income wasn't relevant. People are not aware of the benefits they get added to or subtracted from their income. While we know the data we have are accurate, we are not sure how that affects the way people really react.

Glenn Harrison commented on the difficulty of disentangling child poverty from all of the other stuff. Is it a canary in the coal mine? Or is one of the opinions people hold that it will just go to the parents? He also noted the need to include budget constraints, use a general equilibrium model, and that there is potentially a problem in the short-scale with the hypothetical.

Larry Schweinhart thought about the problem in the context of political traction. For example, he thought of a state representative from his area who is on the appropriations committee that was dealing with a proposal to eliminate all preschool programs in Michigan. He took the vote on the committee panel, and is now running for governor on a plan arguing for universal school. What would a politician need to do to rationalize his position?





Panel VI: Extending shadow prices available for benefit-cost analysis of social programs

In her opening remarks, **Lynn Karoly** noted that much of this conference has been focused on how we can use BCA to evaluate social programs, mainly looking within a given social policy domain and determining whether something is a better investment than another program. We may also want to be able to look *across* domains. Is the state of the art at a point where we can do this? We may not even be there for within a particular social domain. Right now we have a whole range of programs that have been rigorously evaluated, and some of them have BCAs, but even the ones that do have BCAs do not have comparable BCAs.

Karoly is currently working on a study showing social policy domains and their outcomes that could be monetized. When you look at a set of policies or interventions, there is an array of different outcomes that potentially could be monetized in a BCA framework. Many types of interventions are affecting multiple outcomes. The ability to monetize does help you, so you can compare multiple programs, unlike in CEA. One challenge with a whole array of outcomes is that you don't always have ideas of what those values should be.

A recent RAND study shows gaps in the application of BCA to social programs. In some social policy domains, all studies included a BCA, in others, just one out of 6 or 7 included a BCA. Some kinds of interventions have outcomes that are more often evaluated in monetary terms, therefore BCA has become expected. Others are affecting outcomes that don't yet have shadow prices to place on them. This is an area where we can make great contributions.

Mark Cohen spoke about the state of the art with regard to the costs of crime. There have been very few BCAs done in the crime area. Many benefits have not been researched or monetized. There are a few economists who work in the crime area, but some of the more recent WTP studies have been done by criminologists who may or may not know the techniques.

Crime affects quality of life, both for victims—including tangible losses such as medical costs and property loss and intangible losses such as pain and suffering—and for the public. But what are the boundaries of “social cost?” How does the external cost compare to the social cost? Is the cost to society of a \$100 theft \$100? Do we consider the value of the externalities created by a \$100 theft?

Lynn Karoly
RAND Corporation

Rosalie Pacula
RAND Corporation

Paul Heaton
RAND Corporation

Mark Cohen
Resources for the Future





REMOVING BARRIERS

Many qualities are difficult to evaluate or measure, for example, behavioral changes in response to the threat of crime, or generalized anxiety that does not translate into actual avoidance. There may be long-term consequences, such as burglary victims moving because they no longer feel safe in their homes.

On the side of offender costs, there is the opportunity cost of time spent in illegal activity, and the lost productivity of convicted offenders in prison. Imprisoned offenders may also have pain and suffering. Should they be considered “outside of society” as incarcerated persons? Partners and families of offenders will also potentially suffer.

There is also the question of average versus marginal costs. Would more police on the street, better early childhood education, or more at-risk youth programs reduce crime and therefore lead to reduced victimization costs? Do these reduce “fear of crime?” Does the size of the program matter? Should analyses be bottom-up, starting with victim and criminal justice costs, or top-down, starting with people’s WTP to reduce crime? Is it possible to put price on fear? If I break my arm by tripping and falling, is it the same as if someone assaults me and breaks my arm? Currently we are unable to distinguish between the health and criminal impacts.

When incomplete analyses are used for benefits and costs, they can be misleading. For example, Illinois had a budget crunch, and decided to figure out what would happen if they let people out from prison early instead of building a new prison. The analysis concluded it was a good program because it saved money—but they may not have considered all the costs.

Not many people have considered the WTP for crime reduction, Cohen asserts. A primarily top-down approach has been used in the US. Studies in the UK tend to use a QALY approach, which is primarily bottom-up. A question remains whether QALYs are accurate measures for crime, particularly for non-health incomes, or for victimless crimes (e.g., property damage). There are many differences among studies that have been done. Many of these can be attributed to method, including some of the differences discussed above. Intangible costs for pain and suffering were based on jury awards. Top-down and bottom-up approaches can be vastly different. In addition, how do you put a price on fear?

There are many crimes that are underrepresented in the current literature, including fraud, human rights violations, child neglect, arson, terrorism, and others. The field would benefit from working on the policy community to develop meaningful tests, if we can achieve buy-in. We also need more comprehensive quality-of-life





measures in the community, including life satisfaction, and more comprehensive WTP approaches.

Paul Heaton spoke on applying crime cost estimates to calculate the returns to police investments, a practical project that he has been working on at RAND. RAND's Center on Quality Policing is producing a non-technical guide to assist policymakers in understanding how cost of crime research can be used in police personnel investments.

The necessary ingredients for a benefit-cost calculation for police hiring include estimates of the cost of crime, estimates of the change in crime as policing levels change, and estimates of the cost of hiring or firing police. Once there are estimates of the costs of crime, they can be used to figure out how changing the number of police changes the crime rate. This is more challenging than it looks. If you map police presence and crime rates, there is actually a positive relationship—though Heaton is not saying that more police contribute to more crime. Empirically disentangling the causal effect of additional police presence is challenging. Estimates of the costs of hiring or firing additional police clearly depend on the policies considered, and the particular environment the officers would be hired into.

Several studies and real-world situations have given RAND the opportunity to compare costs and benefits of additional police presence. In addition, they have had the opportunity to talk to police chiefs, mayors, and others to see what they believe. Altogether, this points to many areas where research can be improved and more effective policies can be created.

Concerns raised by policymakers include that the tangible costs are credible where the intangible costs are hard to understand. Estimates on the value of a homicide—as much as \$11 million in some studies—are too high and are unbelievable. In part this is due to the fact that contingent valuation studies and survey information is hypothetical and often difficult to explain.

Valuing all crimes equally may also be problematic. For example, some people believed that gang-related homicides should be considered differently from general homicides, based in part on the general identity of the victim. The crimes that have been indexed are nearly all violent crimes. More information needs to be gathered about how to value white-collar, drug, and unreported crimes. How should police functions aside from crime-fighting be valued?

The research that would allow academia to better address policymakers' concerns includes making the tests more convincing, perhaps through using a broader set of





REMOVING BARRIERS

approaches for estimating crime costs. If different methods are used with the same or similar results, this could be more convincing for policymakers. Determining cost estimates for a wider variety of crimes would also grant additional credibility. In addition, as mentioned, we know very little about people's WTP for police officers' other duties, for example, traffic enforcement and disaster response. We also know very little about peoples' WTP and WTA related to racial profiling, excessive use of force, or, for example, their WTA to avoid being stopped.

Lynn Karoly presented an abbreviated version of **Rosalie Liccardo Pacula's** presentation on an alternative framework for estimating the economic cost of drug abuse, because Pacula was unable to be at the conference due to travel delays. Essentially this research is trying to improve on the standard approach of using cost-of-illness (COI) measures for calculating the costs of crime and drugs. A bottom-up approach is typically used that considers, among other things, drug treatment, drug-related morbidity and mortality, productivity losses, crime, other social system costs, and prevention and research. Some of the weaknesses in the current approach can be addressed within the COI framework, while others cannot. For example, COI will not identify which policies are most cost-effective to reduce costs, and ignores chronicity of use and feedback loops.

RAND's alternative framework begins with an epidemiological model of drug use, and then adds costs. This is a microsimulation model that enables researchers to more accurately capture individual heterogeneity, uncertainty in use and outcomes, and feedback loops related to drug problems, without taking away the ability to draw population-based inferences.

Discussion: Extending shadow prices available for benefit-cost analysis of social programs

Bill Cartwright commented on barriers, specifically that the federal government is often impotent when it comes to making policy. The ability to have interaction at the state and local level is something that has been pointed out and should continue to be considered. Cartwright also noted that at the National Institute on Drug Abuse, he had established a research program on prevention and drug abuse. Since he has left, it hasn't backslid. A lack of economists in federal programs is a critical barrier that comes up for all of these programs. Researchers have also found that when they do market results, people do not want to hear about social costs. You have to talk about the taxpayer perspective.





REMOVING BARRIERS

Mark Cohen noted that there was an interesting study in Washington state, where the legislature has mandated that programs be benefit-cost analyzed. Steve Aos is a particular researcher to mention. These studies start with pure taxpayer benefit, and then go to social cost benefits. They highlight the question of taxpayer costs.

Paul Anton noted that one of the challenges is a structural problem that work on crime is done by criminologists. The models themselves don't lend themselves to the sort of structural approaches that would allow you to feed parameters into an empirical model. However, all of this demonstrates that there is a lot of fodder for smart criminologists and economists to do good work.

Scott Farrow commented on the interdomain comparisons. If anyone wants to address these concerns, one of the areas where we do not have much guidance or research is in evaluation across categories. To do these types of comparisons may be risky, but a lot of times they are asked for.

Lynn Karoly agreed that these types of comparisons are risky and are going to happen regardless. Not everyone measures the same outcomes, but the more disclosure we have, the better off we will be. The better information we have, the more able we will be to fill gaps. Though she didn't mean to imply that cross-domain comparisons shouldn't be done, if they are done now, they may not be fruitful.

Mark Cohen noted they were working on a project for PEW that was trying to look across domains. It has been really difficult to put things together, though when you do, interesting things show up. Then the question is what interventions are there? There is value to doing these comparisons.





Panel VII: Barriers to policymakers’ use of rigorous research and benefit-cost findings

Russ Whitehurst spoke about the standard economic model of human behavior and how it may or may not apply to education. The typical model says that the supply side provides evidence of a more effective or efficient research program, followed by wide dissemination of the program where self-interest is served by adopting this more efficient program, based on the information available. This greater demand for more effective programs would drive less effective programs out of the model and out of the market. However, education does not quite work that way. Whitehurst made his case with a particular example, and then argued how things should change.

In education, for example, which has been on the front burner for several administrations, there is a lot of consensus that something should be done, and that we want to do what is best and most effective and efficient. One proposal has been charter schools or school choice more broadly. There are clear positive effects in certain urban areas for certain charter schools. A study that examined 93% of the charter schools found effect sizes of .09 for math and .06 for reading for a year in a charter school. These are not effect sizes to be dismissed, but are relatively small. Expenses for charter schools are high in the startup sense.

Another proposal has been reconstituting the teacher workforce. This draws on research by Coleman in the 1960s. The argument is that a lot of the difference is attributable to differences in classrooms—in short, differences in teachers. As much as 10-15% of the variance has been attributed to teachers. People have then argued that if you could move these teachers from classrooms at the 25th percentile of effectiveness to the 75th percentile, you would create a standard deviation of change in a year. If that were done three years in a row, the racial achievement gap would be eliminated. The fly in the ointment, however, is how to move that child or that teacher once, let alone for three years in a row. Incentives and professional development opportunities are one way. Teach for America has a .15 effect size for math. Merit pay—which has its strongest trial in India, where the incentives were very large compared to the base salary—had effect sizes as high as .15. Though this may be a viable option, it is very expensive.

The Obama Administration has endorsed stronger standards, though there is no known positive effect size for standards. Some states with bad standards have had

Jon Baron
Coalition for Evidence-Based Policy

Russ Whitehurst
Brookings Institution

Daniel Cole
Indiana University School of Law

Michael Greenstone
Massachusetts Institute of Technology

Susan Dudley
Regulatory Studies Center,
The George Washington University





good progress. Standards can definitely be argued for based on other reasons, but from an empirical basis, may not be justified.

Another proposal is curriculum or pedagogy changes. We know a lot about the effect sizes of curriculum, but even then, more research is needed, for example, with comparative effectiveness trials with competing curriculums assigned to schools. Instead of comparing business as usual to the new curriculum, we should compare two new curricula that have the same goals. Some companies have tied up almost the entire curriculum market in America. Math achievement outcomes between the most and least effective curricula were .30. Comparative effectiveness with middle school reading curriculum found none better than the business as usual curriculum. One actually did significantly worse. Most are quite successful, with effect sizes as high as 1.0.

No one is paying attention to curriculum, however. Everyone is talking about implementation, not curriculum. Curriculum is essentially a “free” good, since the cost differences are essentially different. The effect sizes can be larger than any of the other effect sizes. Policymakers should be paying attention to this; what needs to change so that people are looking at evidence like this? One reason why it’s not likely to be considered is that there is no market. Education is effectively a monopoly. Public schools are not competing with private schools, and are not competing with each other. Like all monopolies, education tends towards preservation of the status quo.

In the end, we find that the information provided is confusing, and there are many voices about what works. Many studies have substantial problems with attitudes, biases and aesthetics. This type of challenge can go to the highest levels of government. The Bush Administration continued to support abstinence education even after research showed that it did not work. The Obama Administration has announced strong support for Upward Bound, which has been shown to have no impact on college-going.

Though what you need depends on what you are trying to do, there are two basics that Whitehurst thinks are necessary. First, to have people in the Office of Management and Budget who understand the value of research—which he believes we now have. Second, we need to create a market at the local level for better evidence. Any school that is a recipient of Title I funding should have to justify with a public document why it is purchasing curriculum above a certain cost. Over time, this would create a drive toward evidence-based judgments. At the same time, we need to empower parents to create a market for public schools.





REMOVING BARRIERS

Michael Greenstone spoke on how to move toward a culture of persistent regulatory experimentation and evaluation. His presentation was based on a chapter for a book on the Tobin project. The government regulates many things: the air we breathe, the schools we go to, the electronics we play with and the toys we give our children. Currently, regulations are evaluated before they exist. They're evaluated *ex ante*, and usually they're evaluated, put on books, and stay there forever. We should learn from what has actually happened, and use that to inform our choices.

In Greenstone's ideal world, BCA would be the way things are done. We would measure the benefits, measure the costs, and then figure out the BCA. The current problem is not a failure of analysis, but rather a failure of evidence. The quality of the evidence usually isn't very good.

One challenge is to solve the fundamental problem of causal inference. We need to use real-world data and create counterfactual situations of worlds that don't have regulation. The optimal way to do that is to have regulations randomly assigned. An alternative is quasi-experimental: often regulations are not going to apply to everyone.

Greenstone proposed four "easy" steps toward regulatory reform. First, experiment! Implement regulations to facilitate evaluation, including small-scale trials, or discrete rules for who gets regulated and who does not. The federal government should also learn from the states, both in what they are doing with evaluation, and what is happening between states that have different regulations. Evaluations need to be funded to be able to collect data and publicly release results. Regulations often cost many millions of dollars, but in the current climate, doing a \$10 million evaluation is seen as out of line.

The second step toward regulatory reform is creating a regulatory review board. Self-evaluation is difficult. The Office of Information and Regulatory Affairs (OIRA) plays an important role, but the first responder to whether something is working is the agency that regulates. These are people who manage programs and are devoted to them. They are capable and caring, but sometimes it is hard to see the shortcomings of their own work. The process of evaluation should be done collaboratively, and have more outside influences. An independent review board would evaluate data, and would consist of interested experts, but not people who were directly involved.

The third step would be incorporating automatic sunset and expansion provisions into legislation—pre-specifying dates for the review board to assess a program and expand where analysis is supportive, or modify or remove regulations that fail to





deliver. This may include an automatic sunset for regulations where an analysis of effectiveness is not conducted.

The fourth step is developing and applying a code of ethics. Experimentation must not endanger subjects' safety. Creating greater credibility for BCA is the linchpin of a regulatory system that produces safe and effective regulation.

Daniel Cole spoke on the statutory, political, and methodological barriers to more and better regulatory BCA. Agencies are almost always required to prepare BCAs for major rules, but they are sometimes prohibited from considering them in decisionmaking. Statutory impediments to considering regulatory BCAs are of dubious effect. Arguably, the methodological and political obstacles to considering BCA are more significant than the statutory or legal barriers. This raises an important normative question: Should we ever make decisions without considering BCAs?

Regulatory impact analyses (RIAs) are essentially BCAs, and have been required since 1995 for all major federal agency actions that *might* cost private entities or state, local, or tribal governments more than \$100 million per year. In addition, since the Nixon administration, executive branch agencies have been required to prepare RIAs or BCAs for major regulations.

Statutes rarely prohibit agencies from *considering* the BCAs they prepare, though there are a few exceptions. The Endangered Species Act, for example, requires that the agency make decisions “solely on the basis of the best scientific and commercial data available,” i.e., not cost. Sometimes costs are considered through informal BCA if the statutes otherwise prohibit formal benefit-cost analysis.

Are these statutory impediments ever really justifiable? What is the interaction between legal rights and cost-benefit balancing? Some consider legal rights to be “trumps” that shouldn’t be considered in the analysis. There have been lots of cases where legal rights have been trumped by exigent circumstances. President Lincoln suspended habeas corpus, for example, and Justice Holmes limited the freedom of speech to shout “Fire!” in a crowded theater. But these analyses are not always positive—consider Holmes’ opinion in *Buck v. Bell*, his famous statement that “three generations of imbeciles are enough” was based on an informal BCA. If legal rights are already subject to informal BCAs, why not subject them to formal BCAs, likely with more reliable results.

As Cole noted, there are methodological barriers to proper use of BCA. For example, agencies disagree about subjective BCA assumptions, including valuations





of non-market goods, and what the discount rate should be. Agencies also disagree about how to treat Knightian uncertainty—how to count risk when the probability is unknown. There is also the fallacy of misplaced concreteness, the pathology of “just give me *some* number.”

These methodological disputes over BCA reflect but also engender political conflict. There is also a perception that BCA leads to inequitable regulatory outcomes, or that it can be used and manipulated to either facilitate or impede collective action. There is also a difference between what Cole calls “Economic BCA” and “Political BCA.” Economic BCA identifies policy alternatives, determines foreseeable impacts, calculates values for all of those impacts, calculates the net present benefit for each alternative, and suggests the choice of the alternative with the greatest net benefit or the lowest net cost. Political BCA, in contrast, first starts with an analysis of whether the proposal is inconsistent with the audience’s prior conceptions and then looks for concrete numbers, even if they are unreliable. The benefits and costs are considered with respect to their expected effect on contributions or the budget, and an expected effect on votes may also be considered. Political BCAs also are often simplified to an extent that hides critical complexities. Perhaps political BCA currently matters more than economic BCA. Part of why this conference was convened is to improve economic BCA.

So how can this be done? One, improve methodological consistency, including an independently generated set of “best practices.” The advantages of *formal* BCA also need to be better explained. Increasing the use of regulatory prompt letters will also help. Make it clear that there is no such thing as an incontestable or irrefutable BCA, and that BCA is one of several regulatory decision-making tools, not a panacea. Finally, you have to realize that BCA cannot take the politics out of politics.

Jon Baron continued the discussion on barriers to policymakers’ use of rigorous research and benefit-cost findings. He spoke of the work the Coalition for Evidence-Based Policy does as a nonprofit, nonpartisan organization whose mission is to increase government effectiveness through rigorous evidence about “what works.” The coalition has worked closely with policy makers, Congress, and federal agencies, and has been successful in influencing legislative language, increasing funding for evidence-based evaluations and programs, and raising the level of debate regarding standards of evidence.

An example of their work is with teen pregnancy prevention programs. They were involved in providing input for both Congress and the administration. This involved providing information about a large





REMOVING BARRIERS

multi-state randomized trial of the Carrera program, a youth development program that includes access to reproductive healthcare. It was shown to reduce females' pregnancies/births by 40-50%. It also suggested increased high school completion and college enrollment. This study was a factor in scaling-up evidence-based teen pregnancy prevention focus by the Obama Administration, likely to be enacted in the next few months.

There are several reasons why examples like this do not occur more often. One, most research findings touted as “rigorous evidence” of effectiveness may not be strong enough to guide major program funding decisions. This is because they are often quasi-experimental studies or studies with other key limitations, such as looking to only immediate post-intervention results, or intermediate outcomes. These findings are better viewed as *suggestive* or *promising* evidence. In addition, promising findings are often overturned in subsequent, more definitive randomized controlled tests. In medicine, 50-80% of interventions found promising in Phase II (mostly non-randomized studies) are shown ineffective in Phase III tests (sizable randomized controlled tests). Similar patterns occur in other subject areas. One barrier is that there just aren't enough strong research findings.

Poor reporting of key research findings may also limit policy impact. The way that research findings are presented makes them inaccessible and less persuasive. The effects of interventions are often not reported in “real world” terms that enable readers to gauge their practical importance. Policymakers often worry about hidden assumptions. In addition, presenting equations and coefficients to a policymaker is likely to be ineffective.

In the coalition's experience promoting evidence-based reform, we have found that the power of ideas over vested interests is commonly underestimated. The strength of the evidence matters, as does clarity in how it is reported.

Susan Dudley spoke about benefit-cost analysis from the perspective of regulation. There has been an Executive Order since at least 1981 that requires agencies to understand the benefits and costs of different alternatives, and to try to maximize the benefits to the extent allowed by law. With that—we should be in fat city—but why aren't we?

For one, the Executive Orders do not cover independent agencies. OIRA does not see the financial data for independent agencies such as the Federal Trade Commission. All regulatory authority is delegated to agencies by Congress. The administrator is the person who makes the decision. In some cases, he or she is not allowed to look at cost data, isn't allowed to hear it, and it can't be discussed in front





of him or her. If you look at the statutory mandate, there is no logical limit except the prohibition of BCA on paper. But agencies often are considering costs, they just can't say they are.

Another issue is that one cannot separate legislation from regulation. Legislation and regulation tend to be in response to crises. In these situations, there may be moral objections, similar to what has been discussed earlier. There is a sense that “we know the right answer,” so why do we need to do a BCA? Certain areas—for example, Homeland Security or climate change—are difficult to conduct BCA in, because of a lack of resources, history, or available data.

But if all of this could be swept aside, is BCA the silver bullet? The National Highway Transportation Safety Administration is good at doing BCAs. They often have a lot of data, including actuarial data. However, they routinely intervene in private decisions, under the argument that they can make people better off by limiting their choices. This BCA argues that the private benefits of regulation (for example, saved lives through increased seatbelt use) exceed the private costs of regulation (for example, a personal preference for not wearing a seatbelt, or physical discomfort from awkwardly positioned straps). Similar personal choice issues and potential limitations arise from the greenhouse gas regulations that are currently out for comment.

The bigger challenge, though, is getting agencies to focus on big issues. Agencies are weak in considering multiple alternatives in comparison to the baseline they are measuring against. This is, in part, where the behavioral research comes in. Hopefully Richard Thaler and Cass Sunstein's recent book *Nudge* can help agencies think about what they can use. But this also may be a cause for concern, in that there is a danger in supplanting people's individual preferences with government preferences.

The notion of automatic sunsets, mentioned earlier, are a hard sell. There is no constituency who is in favor of sunsets—not even business. If you're already complying, why would you want your recent-entry competitors to not be bound by the regulations? Clearly, beneficiaries of the regulations will also want them to continue. Ultimately, BCA is a necessary part, but not all of the regulatory analysis. One cannot take the politics out of the discussion, but the more transparent we can make the process, the better off we will be.





Discussion: Barriers to policymakers' use of rigorous research and benefit-cost findings

Trish Hall noted that a review of safe drinking water is required every 6 years. The real problem is lack of data because states are not providing it. Maybe access to data is something to think about when designing regulation. The Office of Management and Budget is currently having trouble with information flexion. If you try to do some experimentation, you get problems with legal issues and human subjects review. It is hard to get through that even if the project only remotely deals with people. Oftentimes it is hard to get expertise in both the substantive area and economics to be on the review board.

Michael Greenstone noted that was exactly his point on designing regulations to allow better data collection. Having been on a scientific advisory board, he knows that internal politics still play a role and limit the questions that can be asked. Perhaps there should be some change in that system as well.

Kerry Smith commented that having BCA requirements changes the design and structure of regulatory programs. Even if only internally, it changes the lens that you are looking through; hopefully it improves it. If we start to think about the regulatory design process and build evaluation on the outside, what will the effect be? Will it change the structure of designing regulations?

Michael Greenstone replied that yes, he believed infusing the culture of rulemaking with BCA from the start would have a tremendous impact in weeding regulations out. In part because you could pin down what the goal of the regulation was. If there is a commitment to a goal, then you can set up the regulation to test whether or not it achieves that goal.

Kim Thompson commented on the ex post issue. One reason to do ex post analysis is because it teaches us about methods and problems with underlying assumptions, as well as reminding us that we need to do a better job thinking about distribution. These impact our ability to improve regulations and target benefits. It also helps promote better communication overall. The community of people who should be asking for post-fact and post-mortem analyses is the agencies themselves.

Susan Dudley asserted that the best example of where ex post analysis has worked was in the analysis of deregulation in the 1980s, including in the work of Fred Kahn and George Stigler, who realized regulations benefitted producers, not consumers.





Appendix A: Conference Program

Unleashing the Power of Social Benefit-Cost Analysis: Removing Barriers

Mon., Oct. 19th

- 7:30 – 8:30 a.m. **Registration / Check-in**
Continental Breakfast, Ballroom III
- 8:30 – 9:00 a.m. **Welcome Remarks**
- 9:00 – 10:30 a.m. **Panel 1: *Barriers to the Use of Benefit-Cost Analysis and Cost-Effectiveness Analysis in Public Health and Medical Decisionmaking***

Chaired by **Mark Pauly**, Bendheim Professor of Health Care Systems, The Wharton School, University of Pennsylvania

John E. Calfee, Resident Scholar, American Enterprise Institute
Peter Neumann, Director, Center for the Evaluation of Value and Risk in Health at the Institute for Clinical Research and Health Policy Studies at Tufts Medical Center; Professor, Tufts University School of Medicine
- 10:30 – 10:45 a.m. **Break**
- 10:45 – 12:15 p.m. **Panel 2: *Barriers to the Valuation of Health Outcomes in Benefit-Cost Analysis and Cost-Effectiveness Analysis***

Chaired by **Lisa Robinson**, Independent Consultant

James Hammitt, Professor of Economics and Decision Sciences, Harvard University, and Director, Center for Risk Analysis
Trudy Ann Cameron, Raymond F. Mikesell Professor of Environmental and Resource Economics, Department of Economics, University of Oregon
Joseph Lipscomb, Professor of Public Health and Georgia Cancer Coalition Distinguished Cancer Scholar, Rollins School of Public Health, Emory University
Dean T. Jamison, Professor of Global Health, University of Washington
- 12:15 – 1:15 p.m. **Lunch**
- 1:15 – 2:45 p.m. **Panel 3: *Behavioral Approaches to Policy***

Chaired by **Michael A. Stegman**, Director of Policy and Housing, John D. and Catherine T. MacArthur Foundation





Unleashing the Power of Social
Benefit-Cost Analysis:

REMOVING BARRIERS

Eldar Shafir, Professor of Psychology and Public Affairs, Princeton University
Hunt Allcott, Research Affiliate, Massachusetts Institute of Technology
Sendhil Mullainathan, Professor of Economics, Harvard University and Managing Director and Founder, Ideas42

2:45 – 3:00 p.m.

Break

3:00 – 4:30 p.m.

Panel 4: *Use of Benefit-Cost Information in Education Policy: Barriers and Opportunities*

Chaired by **V. Kerry Smith**, W.P. Carey School of Business, Arizona State University

(Organized by **W. Steven Barnett**, Board of Governors Professor, Rutgers University and Co-Director, National Institute for Early Education Research)

Clive Belfield, Associate Professor of Economics, Queens College

Tom Lamb, Government Relations Director, PNC Financial Services Group, Inc.

Michael Rebell, Executive Director, Campaign for Educational Equity and

Professor of Law and Educational Practice, Teachers College, Columbia University

Al Lindseth, Of Counsel, Sutherland Asbill & Brennan LLP

6:30 – 9:00 p.m.

Dinner at Zola Restaurant

800 F Street NW

Washington D.C., 20004

Tues., Oct. 20th

7:15 – 8:00 a.m.

Check-in

Continental Breakfast, Ballroom III

8:00 – 9:30 a.m.

Panel 5: *Estimating the Willingness to Pay of the American Public for Childhood Poverty Reduction*

Chaired by **Dale Whittington**, Professor of Environmental Sciences and Engineering, City and Regional Planning, and Public Policy, University of North Carolina at Chapel Hill, and Professor, Manchester Business School (UK)

Co-Chaired by **F. Reed Johnson**, Senior Fellow and Principal Economist, Research Triangle Institute

V. Kerry Smith, W.P. Carey Professor of Economics, Arizona State University

Vic Adamowicz, Professor, University of Alberta

Glenn Harrison, C.V. Starr Chair of Risk Management & Insurance Director, Center for the Economic Analysis of Risk (CEAR), J. Mack Robinson College of Business, Georgia State University





Unleashing the Power of Social
Benefit-Cost Analysis:

REMOVING BARRIERS

9:30 – 11:00 a.m.

Panel 6: *Extending the Shadow Prices Available for Benefit-Cost Analysis of Social Programs*

Chaired by **Lynn Karoly**, Senior Economist, RAND Corporation

Rosalie Pacula, Senior Economist, RAND Corporation

Paul Heaton, Associate Economist, RAND Corporation

Mark Cohen, Vice President for Research, Resources for the Future

11:00 – 11:15

Break

11:15 – 12:45 p.m.

Panel 7: *Barriers to Policymakers' Use of Rigorous Research and Benefit-Cost Findings*

Chaired by **Jon Baron**, President, Coalition for Evidence-Based Policy

Daniel Cole, R. Bruce Townsend Professor of Law, Indiana University -
Indianapolis

Grover "Russ" Whitehurst, Director, Brown Center on Education Policy, The
Brookings Institution

Michael Greenstone, Chief Economist, White House Council of Economic
Advisors

12:45 – 2:00 p.m.

Luncheon, Ballroom II

"A Private Grant Maker's Relentless, Unapologetic Use of Benefit-Cost Analysis"

Special keynote speaker Michael Weinstein

Senior Vice President of the Robin Hood Foundation

Former economics columnist for The New York Times





Appendix B: Conference Participants

Kenneth Acks	Executive Managing Director, The Cost-Benefit Group, LLC
Vic Adamowicz	Professor, University of Alberta
Matthew Adler	Leon Meltzer Professor of Law, University of Pennsylvania Law School
Hunt Allcott	Assistant Professor of Economics, NYU and MIT
Paul Anton	Chief Economist, Wilder Research
Jon Baron	President, Coalition for Evidence-Based Policy
Peter Belenky	Economist, U.S. Department of Transportation
Clive Belfield	Associate Professor, Queens College, City University of New York
Stephen Bell	Principal Scientist & Fellow, Abt Associates, Inc.
Richard Belzer	President, Regulatory Checkbook
Amy Bernstein	Chief, Analytic Studies Branch, Office of Analysis and Epidemiology, CDC/NCHS
Glenn Blomquist	Professor of Economics and Public Policy, University of Kentucky
Samuel Bonti-Ankomah	Principal Analyst, Treasury Board of Canada
Debra Borie-Holtz	Adjunct & Doctoral Candidate, Rutgers University
Carol Bray	Senior Economist, GAO
Brian Bresnahan	Research Assistant Professor, University of Washington
John E. Calfee	Resident Scholar, American Enterprise Institute
Trudy Ann Cameron	Professor, University of Oregon
William Cartwright	Assistant Professor, George Mason University
Michael Caudell-Feagan	Deputy Director, Pew Center on the States
Stephanie Cellini	Assistant Professor, George Washington University
Valerie J. Chang	Program Officer, John D. and Catherine T. MacArthur Foundation
Inimai Chettiar	Legal Fellow, Institute for Policy Integrity
Mark Cohen	Vice President for Research, Resources for the Future
Daniel Cole	R. Bruce Townsend Professor of Law, Indiana University School of Law - Indianapolis
Joseph Cook	Assistant Professor, Evans School, University of Washington
Joseph Cordes	Prof. of Economics, Public Policy and Public Administration, and International Affairs, Trachtenberg School of Public Policy and Public Administration
Mark Courtney	Professor, University of Washington
Dennis Culhane	Professor, University of Pennsylvania
Tyler Davis	PhD Student, University of Washington
Victoria Deitch	Technical Research Analyst, MDRC
Libby Doggett	Deputy Director, Pew Center on the States
Susan Dudley	Director, Regulatory Studies Center, George Washington University
John Fantuzzo	Albert M. Greenfield Professor Human Relations, University of Pennsylvania
Scott Farrow	Professor and Chair, UMBC/Department of Economics
Lou Garrison	Professor, University of Washington
Michael Greenstone	Professor of Economics, Dept. of Economics, MIT
Fumie Yokota Griego	Economist, Office of Management and Budget, Office of Information and Regulatory Affairs
Trish Hall	Economist, US EPA
James K. Hammitt	Professor of Economics & Decision Sciences, Harvard University
Michael Hanemann	Chancellor's Professor, University of California, Berkeley
Kevin Haninger	AAAS Fellow, EPA
Arnold Harberger	Professor of Economics, UCLA
Glenn Harrison	Director, Center for the Economic Analysis of Risk, Georgia State University
Paul Heaton	Associate Economist, RAND Corp.
Monica Herk	Consultant
Michael Jacobson	President/Director, Vera Institute of Justice
Dean Jamison	Professor of Global Health, Department of Global Health, University of Washington





Unleashing the Power of Social
Benefit-Cost Analysis:

REMOVING BARRIERS

Glenn Jenkins	Professor, Queen's University and Eastern Mediterranean University
Amber Jessup	Economist, DHHS/ASPE
F. Reed Johnson	Senior Fellow, Research Triangle Institute
Lynn A. Karoly	Senior Economist, RAND Corporation
Mark Kipperman	Senior Cost Analyst/Systems Engineer, BAE Systems, Inc.
Allen Klaiber	Assistant Professor, Pennsylvania State University
Mary Kokoski	Managing Editor, Journal of Benefit-Cost Analysis
Jamie Kruse	Director, Center Natural Hazard Research, East Carolina University
Kerry Krutilla	Associate Professor, Indiana University Bloomington
Thomas Lamb	Senior Vice President, PNC Financial Services Group
Cynthia Lamy	Metrics Manager, Robin Hood
Alfred Lindseth	Of Counsel, Sutherland
Joseph Lipscomb	Professor of Public Health, Emory University
Mark Long	Associate Professor, University of Washington
Kevin Marsh	Head of Economic Evaluation, Matrix Knowledge Group
Cristina McLaughlin	Economist, FDA CFSAN
Greg Michaels	Director, Impact Measurement, LAC, Chemonics International
Seong Hyeok Moon	PhD Candidate, University of Chicago
Manuel Moreno	Director of Research, Chief Executive Office LA County
John Morrall	Consultant
Sendhil Mullainathan	Professor of Economics, and Managing Director and Founder, ideas42, Harvard University, Department of Economics
Jon Neill	Professor, Western Michigan University
Peter Neumann	Director, Center for the Evaluation of Value and Risk in Health, Tufts Medical Center
Susan Offutt	Chief Economist, GAO
Semra Ozdemir	PhD Student, University of North Carolina at Chapel Hill
Richard Peck	Associate Professor, University of Illinois at Chicago
Peter Pecora	Senior Director Research/Faculty, Casey Family Programs/University of Washington
Michael Rebell	Executive Director, Campaign for Educational Equity
Justin Robertson	Graduate Student, University of Washington
Lisa Robinson	Independent Consultant
Dan Rosenbaum	Senior Economist, Economic Policy, Office of Management and Budget
Kyle Ross	Senior Fellow, PORPP, University of Washington
Robert Scharff	Assistant Professor, Ohio State University
Andrew Schmitz	Eminent Scholar, University of Florida/Food and Resource Economics Department
Jason Schwartz	Legal Fellow, Institute for Policy Integrity
Larry Schweinhart	President, HighScope Educational Research Foundation
Eldar Shafir	Professor of Psychology and Public Affairs, Princeton University
David Smallwood	FANRP Director, USDA-ERS
V. Kerry Smith	Professor, Arizona State University
Ardith Spence	Senior Economist, U.S. Government Accountability Office
Michael A. Stegman	Director of Policy and Housing, John D. and Catherine T. MacArthur Foundation
Judy Temple	Associate Professor of Public Affairs, University of Minnesota
Kimberly Thompson	President, Kid Risk, Inc.
Craig Thornton	Managing Director for Health Research, Mathematica Policy Research
David Vanness	Assistant Professor, University of Wisconsin
Sara Watson	Senior Officer, The Pew Charitable Trusts
Michael Weinstein	Chief Program Officer, Robin Hood Foundation
Barry White	Research Fellow, University of Minnesota
Russ Whitehurst	Director, Brown Center on Education Policy, The Brookings Institution
Dale Whittington	Professor, University of North Carolina at Chapel Hill
Franck Wiebe	Chief Economist, Millennium Challenge Corporation
Richard Williams	Director, Regulatory Studies and Government Accountability, Mercatus Center at GMU
Richard Zerbe	D. J. Evans Distinguished Professor, The Evans School, University of Washington
David Zorn	Director, Division of Social Sciences, FDA

