



**World Health
Organization**

REGIONAL OFFICE FOR **Europe**



**Don Kenkel
Marc Suhrcke**

ECONOMIC EVALUATION OF THE SOCIAL DETERMINANTS OF HEALTH

**An overview of conceptual and
practical issues**



**World Health
Organization**

REGIONAL OFFICE FOR **Europe**

ECONOMIC EVALUATION OF THE SOCIAL DETERMINANTS OF HEALTH

An overview of conceptual and practical issues

Don Kenkel, PhD

Department of Policy Analysis and Management
Cornell University
Ithaca NY
USA

Marc Suhrcke, PhD

School of Medicine, Health Policy and Practice
University of East Anglia
Norwich
United Kingdom

Keywords

COST-BENEFIT ANALYSIS
PROGRAM EVALUATION
HEALTH ECONOMICS
SOCIOECONOMIC FACTORS
HEALTHCARE DISPARITIES
HEALTH STATUS DISPARITIES
EUROPE

Address requests about publications of the WHO Regional Office for Europe to:

Publications
WHO Regional Office for Europe
Scherfigsvej 8
DK-2100 Copenhagen Ø, Denmark

Alternatively, complete an online request form for documentation, health information, or for permission to quote or translate, on the Regional Office web site (<http://www.euro.who.int/pubrequest>).

© World Health Organization 2011

All rights reserved. The Regional Office for Europe of the World Health Organization welcomes requests for permission to reproduce or translate its publications, in part or in full.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either express or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use. The views expressed by authors, editors, or expert groups do not necessarily represent the decisions or the stated policy of the World Health Organization.

Cover art: Deanna Sirlin, *After Me #6*, 2009, 16 x 8 inches, mixed media

CONTENTS

	<i>Page</i>
Acknowledgements	6
Executive summary	7
Introduction	8
1. Outline and scope of this report.....	12
2. Measuring the effectiveness/causal impacts of SDH interventions	16
2.1 Identifying the impacts of SDH interventions	16
2.2 Experimental methods to estimate the causal impacts of SDH interventions	17
2.3 Quasi-experimental econometric methods to estimate the causal impacts of SDH interventions	19
2.4 Structural econometric methods to estimate the impacts of SDH interventions	22
2.5 Recommendations	24
3. Valuing the consequences of SDH interventions	28
3.1 Valuing costs	28
3.2 Cost-Effectiveness and Cost-Utility Analysis	29
3.3 Cost-Benefit Analysis.....	30
3.4 Recommendations	33
4. Valuing reductions in health inequities	36
4.1 The equity-efficiency trade-off.....	36
4.2 Valuing reductions in health inequities in Cost-Effectiveness Analysis	38
4.3 Valuing reductions in health inequities in Cost-Benefit Analysis	40
4.4 Recommendations	41
References.....	46

Acknowledgements

Draft prepared for the World Health Organization, Regional Office for Europe. Chris Brown of the WHO European Office for Investment for Health and Development has provided useful guidance and comments. J. Catherine Maclean and Kerri Lavelle, Cornell University, have provided excellent research assistance. The authors are solely responsible for all views contained in this report, and for its remaining gaps and limitations.

Executive summary

This Report develops a framework for the economic evaluation of social determinants of health (SDH) interventions. The use of economic arguments, in particular regarding the “value for money” of suggested interventions, has so far been a low priority in recent major SDH initiatives, such as in the WHO Commission on Social Determinants of Health. At the same time, the need to add an economic perspective to the analysis of SDH and of health inequalities has been increasingly recognized in the public health community.

Because any economic evaluation hinges on the evidence of effectiveness in the first place, a considerable share of our discussion focuses on the challenge of assessing whether a given intervention “works” (and if so, for whom it does). We stress the importance of using a research design that provides credible estimates of the causal impacts of the intervention under consideration, even in the absence of randomised experimental evidence. The approaches we propose, using quasi-experimental, econometric or structural models, can be used to provide credible estimates of the effectiveness of SDH interventions. In a framework for economic evaluation, the choice between these methods will usually depend on the existing research base and the practicality of new research on the causal impacts of the SDH intervention being evaluated.

For the purpose of valuing the health improvements and indeed other non-health effects possibly resulting from SDH interventions, we recommend social cost-benefit analysis as the approach to develop a comprehensive measure that reflects the value of improving outcomes across multiple domains including health, earnings, and crime. We focus on the application of cost-benefit approaches to economic evaluations of SDH interventions.

Last not least we discuss the implications of SDH interventions for health equity, and how distributional consequences might be taken into account in an economic evaluation. While there is growing acceptance among economists for the need to capture and take into account distributional consequences along-side economic evaluations, and while we also know that people are in principle willing to sacrifice overall health benefits for a reduction in health inequalities, a universally accepted method to incorporate the value of reducing health inequities into economic evaluations has yet to emerge. Instead, we may think of a hierarchy of approaches to incorporating equity considerations into economic evaluations of SDH interventions.

Introduction

Policy-makers, on behalf of the public they represent, face difficult choices about allocating scarce societal resources to investments that address the social determinants of health (SDH). The WHO Commission on the Social Determinants of Health (CSDH) describes the SDH as including the conditions of daily life “in which people are born, grow, live, work and age, including the health system.” Specific interventions to improve the conditions of daily life include investments in early childhood education, environmental quality, health care, housing, nutrition, schooling, and work conditions.

This report develops a framework for the economic evaluation of SDH interventions. The use of economic arguments, in particular regarding the “value for money” of suggested interventions, has so far been a low priority in recent major SDH initiatives, such as in the WHO CSDH (Epstein et al, 2009).¹ At the same time, the need to add an economic perspective to the analysis of SDH and of health inequalities has been increasingly recognized in the public health community (Lavin & Metcalfe 2009).

Because any economic evaluation hinges on the evidence of effectiveness in the first place, a considerable share of our discussion focuses on the challenge of assessing, whether a given intervention “works” (and if so, for whom it does). It is well known that compared to clinical interventions, the evidence base for broader, population-based public health interventions (and this may include at least some SDH interventions) is much thinner (Petticrew et al 2005; Wanless 2004). The availability of abundant evidence on the inverse relationship between socioeconomic status and health can be seen as a useful target and benchmark for public policy, but it does not inform policy-makers about the best course of action to take.

Once effectiveness is reasonably well established, economic evaluation methods can be applied to provide systematic guidance to policy makers about the costs and consequences of alternative courses of action. Frequently, economic evaluations are limited to sector-specific domains. For example, the costs of a new medical treatment will be compared to the life years gained, or the costs of worker training will be compared to the benefits of higher lifetime earnings. In contrast, economic evaluations of SDH interventions must take into account their impacts across multiple domains. For example, across the

¹ There was some consideration of certain economic aspects in the recent England-specific Strategic Review of Health Inequalities in England post 2010. (The Marmot Review, 2010), in that there was an attempt to estimate the expected economic benefits of reducing health inequalities (Mazzucco, Meggiolari and Suhrcke, 2010). This exercise could, however, not take into account any specific interventions or their associated costs. For further economic background work to the Marmot Review, including the background papers of the Economics Task Group, see <http://www.instituteofhealthequity.org/projects/fair-society-healthy-lives-the-marmot-review>.

world investments in nutrition and early childhood education have shown potential not only to improve health, but also to lead to cognitive and behavioural gains and increased schooling (Nores and Barnett 2010). Economic evaluations of SDH interventions must also recognize the value society places on reductions in health inequities. An additional criterion is for greater transparency in evaluation methods.

1. Outline and scope of this report

1. Outline and scope of this report

As defined in a standard textbook: “the basic tasks of any economic evaluation are to identify, measure, value and compare the costs and consequences of the alternatives being considered.” (Drummond et al. 1997, pp. 8–9). Section 2 of this report discusses identifying and measuring the consequences of SDH interventions. Section 2 stresses the importance of using a research design that provides credible estimates of the causal impacts of the intervention under consideration. Section 3 discusses valuing the health improvements due to SDH interventions. Section 3 proposes social cost-benefit analysis as the recommended approach to develop a comprehensive measure that reflects the value of improving outcomes across multiple domains including health, earnings, and crime. Standard references provide in-depth discussions of the general methods of programme evaluation (Cook and Campbell 1979, Angrist and Pischke 2009), social cost-benefit analysis (Boardman et al. 2006), and cost-effectiveness analysis (Gold et al. 1996, Drummond et al 1997). The aim of sections 2 and 3 is not to provide a complete review of that material, but to focus on the problems of applying those methods to conduct economic evaluations of SDH interventions.

Economists evaluate policy interventions in terms of their economic efficiency and their implications for equity. The methods of cost-effectiveness analysis and social cost-benefit analysis focus on the economic efficiency of interventions. Section 4 of this report discusses the implications of SDH interventions for equity, especially their potential role in reducing health inequities. While there is growing acceptance among economists for the need to capture and take into account distributional consequences along-side economic evaluations, no universally accepted method to incorporate the value of reducing health inequities into economic evaluations has yet emerged.

The last sub-section of sections 2, 3, and 4 summarize the recommended framework for economic evaluations of SDH interventions.

The WHO CSDH (2008) describes three principles of action that guide its policy recommendations about the social determinants of health:

- “1. Improve the conditions of daily life - the circumstances in which people are born, grow, live, work, and age.
2. Tackle the inequitable distribution of power, money, and resources - the structural drivers of those conditions of daily life - globally, nationally, and locally.

3. Measure the problem, evaluate action, expand the knowledge base, develop a workforce that is trained in the social determinants of health, and raise public awareness about the social determinants of health.”

The methods discussed in sections 2 and 3 are best-suited to evaluate specific SDH interventions to improve the conditions of daily life. The focus on specific interventions is consistent with the first principle of action, and the focus on evaluation is consistent with the third principle of action. As discussed in section 4, specific SDH interventions also have the potential to reduce some of the inequities referred to in the CSDH’s second principle. However, the scope of this report does not cover the second principle’s broader focus on the structural drivers of the distribution of power, money, and resources.

2. Measuring the effectiveness/causal impacts of SDH interventions

2. Measuring the effectiveness/causal impacts of SDH interventions

2.1 Identifying the impacts of SDH Interventions

The first steps of an economic evaluation are also part of standard programme evaluation: identify and measure the consequences of the intervention under consideration. Examples of SDH interventions include investments in nutrition, early childhood education, housing, schooling, work conditions, housing, health care, and the quality of the natural environment. By definition of SDH, these interventions potentially have important health impacts. Beyond the health impacts, identifying potential consequences will vary across examples. For example, Karoly et al. (2001) discuss how to measure the benefits of early childhood education programs. They present four broad domains of programme impacts or consequences: emotional and cognitive development; education; public assistance receipt, income, crime; and health. Within broad domains like these, the next step is to develop quantifiable measures. Karoly et al.'s list of illustrative measures within the domain of emotional and cognitive development includes: socio-emotional and behavioural scores; IQ test scores; and teacher's ratings. In another example, Carlson et al (2011) consider a housing voucher programme for low-income families in the United States (US). In addition to improved housing quality and neighbourhood environment, Carlson et al identify increased child achievement, as measured by years of schooling and reduced criminal behaviour, as potential impacts of the housing voucher programme.

Economic evaluations consider a broad range of consequences and should include the intervention's direct impacts on programme participants and the indirect impacts on others in society. Indeed, the societal impacts, or in economics terminology the positive externalities, provide the economic efficiency rationale for public sector intervention.² In addition to the private returns enjoyed by the participants, SDH interventions may also yield societal returns, for example if participants are less likely to require public assistance, or are less likely to commit crimes. Measuring these impacts poses special challenges. It will often require additional information, beyond administrative records and measures routinely collected for non-economic programme evaluation. In addition, the societal returns to an SDH intervention may only accrue years after the programme, for example when participants in an early childhood education programme commit fewer crimes as adolescents. When

² It is standard in economics to make the conceptual distinction between economic efficiency and equity rationales for public sector interventions. As will be discussed in more detail below, reductions in health and other inequities also provide much of the rationale for SDH interventions.

these impacts are not measured and quantified, the economic evaluation should still attempt to identify all potential impacts and discuss the implications of the omissions. The goal is to provide as much information as possible about how the measured benefits of the programme compare to the total benefits.

2.2 Experimental methods to estimate the causal impacts of SDH interventions

The most challenging task in this step of an evaluation is to develop credible estimates of the causal impacts of the SDH intervention under consideration. The introduction mentioned the widely observed and well-documented inverse associations between socioeconomic status and health. However, it is also widely observed that associations do not prove causation, so this evidence does not establish whether SDH interventions that address socioeconomic disparities will necessarily cause improvements in health and reductions in health disparities.³

The effectiveness or causal impact of an intervention can be defined by the comparison of observed outcomes with the intervention to the counterfactual outcomes that would have resulted in the absence of the intervention. For example, by this definition the effectiveness of an early childhood intervention on intelligence quotient (IQ) test scores is the difference between a participant's score with the intervention and what his or her score would have been without the intervention. The hypothetical counter-factual is unobservable. Experimental, quasi-experimental, and econometric methods are used to develop estimates of interventions' effectiveness or causal impacts.

Randomized experiments are often viewed as the "gold standard" for estimating effectiveness. The random assignment of cases to intervention and control groups provides a strong basis for estimating causal treatment effects. Randomized clinical trials (RCTs) are the mainstay of much biomedical research. For example, the prescription drug approval process typically rests on RCTs that compare the health outcomes for patients given a new pharmaceutical product to patients in the control group. RCTs are also used to estimate the effectiveness of public health interventions. In a review of research published between 2000 and 2005, Weatherly et al (2009) identify 154 studies that conduct economic evaluations of

³ To mention but one example from obesity context, a recent literature review (Saelens, Sallis and Frank 2003) suggests a significant link between characteristics of the built environment and the prevalence of obesity. It finds that those who live in close proximity to fast-food restaurants, or have little opportunity to undertake physical recreational activities, tend to be at a higher risk of obesity. Such studies are useful for identifying the types of people who live near to particular facilities but are unlikely to be sufficient in terms of demonstrating a link between the existence of those facilities and obesity levels. The observed correlation could also arise because those who like to engage in physical activity decide to locate preferably in the neighbourhood of recreational activities ("reverse causality") or because there are third, unobserved factors that affect both obesity and people's proximity to an "obesogenic" built environment.

interventions in eleven areas of public health: accidents; alcohol; ante-natal and post-natal visiting; drug use; HIV/AIDS; low birth weight; obesity and physical activity; sexually transmitted infections; smoking; teenage pregnancy; and youth suicide prevention. Fifty-eight of these studies were based primarily on an RCT.

By contrast, randomized experiments have only very rarely been conducted to evaluate the broader interventions and policies that address the SDH. Notable US examples include experimental evaluations of early childhood interventions such as the Perry Pre-School project, and an experimental evaluation of the Moving to Opportunity experiment that provided housing vouchers to poor families. Outside the US, a recent meta-analysis of evaluations of early childhood interventions found 23 estimates of effectiveness based on randomized designs (Nores and Barnett 2010).⁴ Despite these examples, a range of political, ethical and practical concerns limit the use of randomized designs to evaluation SDH interventions. In the UK, for instance, there was no randomization of the communities that had access to the major early childhood programme “Sure Start” (loosely modelled on the US “Head Start” programme), because doing so was considered politically unacceptable (MacIntyre 2003).

Despite the challenges to implementing experimental designs in SDH interventions, it is important to bear in mind that at least some of the concerns can be overcome. For instance, instead of completely “depriving” one or more communities from the “treatment”, all communities could receive the intervention, but in a phased manner, thereby allowing analysis of the variation in outcomes according to the intensity of the intervention over time (Jones 2006).

On the other hand, the recognition of the RCT as the gold standard should not ignore its limitations, arguably the main one being the very restricted generalisability (external validity) of the findings beyond the population and circumstances encountered within the trial. The RCT is a strong research design in terms of the internal validity of its answer to the research question: Did the intervention truly cause changes in outcomes? For evaluating SDH policies, it will often also be crucial to address external validity: Can the results be extrapolated or generalized beyond the specific experiment studied? In the context of clinical RCTs, the issue of external validity is related to the distinction between efficacy and effectiveness. RCTs demonstrate the efficacy of a treatment under tightly controlled, often nearly ideal, conditions. Demonstrated efficacy does not always translate into effectiveness of treatments as they are delivered and used in the real world. Because of the complicated

⁴ For a critical review of the evaluative literature of the UK’s “Sure Start” programme and for coverage of some other European literature on early childhood interventions, see Pillas and Suhrcke (2009).

nature of many SDH interventions, there is even more scope for the effectiveness of an intervention in a new context (for example, a different country or time) to be much different than the efficacy estimated in an experimental evaluation. Beyond that, in many or most circumstances new SDH interventions being considered will not be exactly like interventions that have been previously evaluated through an RCT, especially because experimental evaluations of SDH interventions are so rare.

Another practical limitation is that an RCT may be unfeasible in cases where the health effects are likely to materialise only far ahead in the future, as is probably the case in many SDH interventions. In addition to their short-term impacts on programme participants, many SDH interventions have potentially important impacts on participants and society over the much longer term. For example, an evaluation of the Perry Preschool Program, an early childhood intervention, examines its impact on lifetime earnings and criminal activity (Heckman et al 2010). An advantage of this study is that programme participants, who were aged 3 when the intervention was delivered, were followed through the age of 40. Still, to complete the evaluation Heckman et al are forced to extrapolate out-of-sample to obtain lifetime earnings profiles. Many evaluations involve much shorter follow-up periods. Furthermore, there is an inherent dilemma: it is impossible to estimate long-term impacts for recent interventions. This results in the frustrating situation of having estimates of the long-term impacts of an intervention launched in the early 1960s (the Perry Preschool Program), while lacking estimates of the long-term impacts of more recent interventions.

2.3 Quasi-experimental econometric methods to estimate the causal impacts of SDH interventions

Quasi-experimental and econometric methods to estimate an intervention's effectiveness rely on finding, in observational data, a 'natural' quasi-experiment that mimics the randomization of a controlled experiment. Studies that use what Rosenzweig and Wolpin (2000) call truly "natural" (as opposed to man-made) natural experiments rely on variation created by natural events such as weather or the birth of twins to estimate causal impacts. Several recent studies use quasi-experimental variation in school assignments created by the hurricane Katrina and the man-made process used to re-locate former New Orleans residents (Imberman, Kugler, and Sacerdote 2009). These examples show how this approach can shed light on the effectiveness of SDH interventions, as long as the natural events are comparable to an intervention of interest. In the Katrina example, the hurricane and its aftermath resulted in changes that could have been achieved through plausible interventions to invest in school quality for disadvantaged youth.

Quasi-experimental methods often rely on man-made quasi-experiments to estimate causal impacts. A prominent class of studies that take this approach are known in economics as using difference-in-difference estimators. As Meyer (1995) explains, in the older programme evaluation tradition this research design is known as the “before and after design with an untreated comparison group.” A typical example in the US context is when a new policy, such as welfare reform, is enacted in some but not all states. The states with the policy form the treatment group, and the other states make up the untreated control group. The first difference is the before-and-after comparison of outcomes in the treatment group of states that adopted the policy. To help rule out confounding from other trends or shocks, this difference is compared to the before-and-after comparison of outcomes in the control states. The difference-in-difference thus provides an estimate of the causal impact of the policy change. Angrist and Pischke (2010) note that the difference-in-difference method is “probably the most widely applicable” of what they call the “design-based estimators” and what others have called the natural experiment approach.

As Meyers (1995) points out, the difference-in-difference approach relies on variation across states which is neither natural nor an experiment. The semantics call attention to the key assumption of the method: that policy differences across states are sufficiently random that they are “like” true experiments. Policy endogeneity thus poses a fundamental challenge to this approach. Policies are not randomly distributed across states. Depending upon the specifics of the political process, the non-randomness might bias estimates either away from or towards zero effect.

Another quasi-experimental method with a longer history in programme evaluation and more recently in econometrics is the method of regression discontinuity. This approach relies on variation that creates a discontinuity in assignment to an intervention. Individuals on each side of the discontinuity are very similar except for their receipt of the intervention, so they provide useful treatment versus control groups. For example, Angrist and Lavy (1999) exploit a discontinuity created by an Israeli school policy that capped class sizes at 40. A cohort of size 41 was usually split into 2 smaller classes of 20 and 21, while a cohort of size 39 remained in a large class. Comparisons of the cohorts who experienced different class sizes allow Angrist and Lavy to estimate the impact of class size on outcomes. Lee and Lemieux (2010, p.350) note that the transparent identification strategy of the regression discontinuity approach typically is “accompanied by an equally transparent graph showing the empirical relation between the outcome and the assignment variable.” While Lee and Lemieux stress the need for regression analysis and formal statistical tests, the transparency of the regression discontinuity approach is an important strength. Lee and Lemieux (2010,

Table 5) list over 50 studies that use the regression discontinuity approach in applications broadly relevant to SDH interventions, including interventions in education, the labour market, health, crime, the environment, and anti-poverty programs.

The econometric method of instrumental variables (IVs) can also be viewed as a quasi-experimental method. A suitable IV provides an exogenous source of variation that allows the econometrician to identify a causal relationship. For example, school reforms in various countries create exogenous variation in years of completed schooling. Economists have used the IV approach in this context to estimate the causal impact of an additional year of schooling on earnings, health, and other outcomes.

Because quasi-experimental evaluation methods use the same logic as randomized experiments, it is not surprising that they tend to share the same limitations regarding generalizability or external validity. As with RCTs, the details and the context of a new intervention being evaluated may differ from the past quasi-experiment studied. For example, a difference-in-difference study of state welfare reforms in the early 1990s might provide a credible estimate of their causal impact on labour market participation. However, the estimate will not necessarily generalize to predict the impact of a welfare reform with somewhat different components or even of an identical welfare reform that is enacted during a time when the macro-economy is not as strong as it was during much of the 1990s in the US.

The econometric method of instrumental variables also yields estimates that may be of limited generalizability. Critics of many applications of the instrumental variables method argue that the “policy questions being addressed are not always clearly stated” (Heckman and Vytlacil 2005, p.670) and question whether the parameter being estimated is “what we want to know” (Deaton 2010, p. 430). For example, the instrumental variables method provides an estimate of the average returns to schooling for a certain sub-population, but an SDH intervention that improved schooling for another sub-population might yield much different returns. As with experimental and other quasi-experimental approaches, the problem is that the quasi-experiment implicit in the instrumental variables method might not be very similar to the new policy being considered. To address this limitation, Heckman and Vytlacil (2005), Deaton (2010) and others advocate a structural econometric approach that emphasizes clearly articulated economic models. The application of the structural approach to evaluating SDH interventions is discussed in more detail in the following section.

In response to the criticisms, many applied econometricians agree that the instrumental variables method can be useful in such applications as long as its limitations are recognized. It is especially important to recognize that the method can yield multiple,

valid estimates of causal treatment effects, for example, multiple, valid estimates of the marginal returns to education for different sub-populations. In this example, the instrumental variables estimate is a weighted average of the causal effect of a year of education within a subgroup, where the weights depend on how much the subgroup is affected by the instrumental variable. As a result, the approach provides estimates of a local average treatment effect (LATE) (Angrist et al 1996). The LATE is “the average treatment effect for those who change treatment status because they comply with the assignment-to-treatment mechanism implied by the instrument.” (Ichino and Winter-Ebmer 1999). With heterogeneous treatment effects, the instrumental variables method does not necessarily provide estimates of the average marginal return to education. But as Card (2001) emphasizes: “For policy evaluation purposes, however, the average marginal return to schooling in the population may be less relevant than the average return for the group who will be impacted by a proposed reform. In such cases, the best available evidence may be IV estimates of the return to schooling based on similar earlier reforms.” More generally, an instrumental variables estimate of the effectiveness of earlier SDH interventions can provide useful evidence about the effectiveness of a similar, new SDH intervention being evaluated.

2.4 Structural econometric methods to estimate the impacts of SDH interventions

As mentioned above, the structural econometric approach is an alternative to quasi-experimental approaches to estimate the causal impacts of SDH interventions. Heckman and Vytacil (2005, p. 669) summarize it as a very broad approach for empirical economics: “Evaluating the impact of public policies, forecasting their effects in new environments, and predicting the effects of policies never tried are three central tasks of economics.... The structural approach emphasizes clearly articulated economic models that can be used to accomplish all three tasks....” In an SDH-related example of this approach, Cunha and Heckman (2009) develop a structural model of capability formation that can be used to evaluate early childhood education interventions. Drawing on research from psychology and human development, the model assumes that children’s cognitive and non-cognitive capabilities are influenced by family environments and other investments. Heckman et al (2008) use this approach to study the Perry Preschool Program’s effects on children’s capabilities that ultimately determine outcomes such as lifetime earnings.

One of the strengths of the method of structural econometrics is that it places special emphasis on the problem of external validity (Heckman and Vytacil 2005). As Nevo (2010) explains, “structural modelling can provide a way to extrapolate observed responses to

environmental changes to predict responses to other not-yet-observed changes.” Instead of a simple rule like linear extrapolation, or the use of expert judgement, structural modelling uses guidance from theory for its extrapolations: “Structural analysis gives us a way to relate observations of responses to changes in the past to predict the responses to different changes in the future. It does so in two basic steps. First, it matches observed past behaviour with a theoretical model to recover fundamental parameters such as preferences and technology. Then, the theoretical model is used to predict the responses to possible environmental changes, including those that have never happened before, under the assumption that the parameters are unchanged.” (Nevo 2010). Cunha and Heckman (2009, p. 44) further argue that structural models provide a way to synthesize evidence on intervention effectiveness from studies of diverse programs and diverse populations. But as Angrist and Pischke (2010) caution: “Empirical evidence on any given causal effect is always local, derived from a particular time, place, and research design. Invocation of a superficially general structural framework does not make the underlying variation or setting more representative. Economic theory often suggests general principles, but extrapolation of causal effects to new settings is always speculative.”

Not surprisingly, the generalizability of the structural approach carries with it corresponding weaknesses stemming from its assumptions and complexity. On the technical side, estimates from structural models can be sensitive to assumptions about functional forms and the statistical distributions of unobservables. In addition, as Heckman (2010, p. 358) notes, “the often complex computational methods that are required to implement this approach make it less transparent” and make replication and sensitivity analyses more difficult. The source of identification in structural models is also often much less transparent than in the quasi-experimental research-design based approach. As another advocate of the structural approach explains: “Here, the phrase ‘how a parameter is identified’ refers instead to a more intuitive notion that can be roughly phrased as follows: What are the key features of the data, or the key sources of (assumed) exogenous variation in the data, or the key a priori theoretical or statistical assumptions imposed in the estimation, that drive the quantitative values of the parameter estimates, and strongly influence the substantive conclusions drawn from the estimation exercise?... In my view, there is much validity to [this] criticism of structural work here. The main positive contribution of the “experimentalist” school has been to enhance the attention that empirical researchers pay to identification in the more intuitive sense noted above.”

A complete discussion of the relative merits of structural econometric models versus quasi-experimental/design-based approaches is beyond the scope of the brief overview in

this report. The merits are discussed in recent exchanges in two of the journals of the American Economic Association: the Spring 2010 issue of the *Journal of Economic Perspectives* and the June 2010 issue of the *Journal of Economic Literature*. Although there are clearly differences of opinion within the economics profession, it is important not to exaggerate the extent of disagreement. Advocates of each approach generally recognize the validity of many of the criticisms from the “other side” of the debate. Moreover, Chetty (2009), Heckman (2010) and other economists are actively involved in “building bridges” between the two approaches. For example, Heckman (2010, p. 359) suggests empirical economics should follow what he calls Marschack’s Maxim: “economists should solve well-posed economic problems with minimal assumptions.”⁵ Heckman explains that for many policy analyses, it is often possible to answer well-posed questions without estimating the individual structural parameters emphasized in the structural approach. For example, it may be possible to answer well-posed questions about the causal impacts of the Perry Preschool Program or other SDH interventions on child outcomes, without estimating the parameters of the structural model of capabilities developed by Cunha and Heckman (2009).⁶ As a bridge between the approaches, “[t]his approach advocates transparency and empirical robustness as does the program evaluation approach, but it also focuses attention on answering clearly stated economic and policy questions.” (Heckman 2010, p. 359)

2.5 Recommendations

The approaches reviewed above can be used to provide credible estimates of the effectiveness of SDH interventions. In a framework for economic evaluation, the choice between these methods will usually depend on the existing research base and the practicality of new research on the causal impacts of the SDH intervention being evaluated. Randomized experimental designs are often seen as the gold standard, but they have limitations, especially regarding generalizability.⁷ Moreover, the existing research base of randomized evaluations of SDH interventions is thin, and new research using this method will often be impractical. In practice, evaluations of SDH interventions will most often rely on quasi-experimental methods, including the econometric method of instrumental variables, to estimate the causal impacts required to measure programme effectiveness.

⁵ The maxim is named after Jacob Marschack, based on an important paper he wrote in the early 1950s during the first wave of structural econometrics.

⁶ Heckman does not suggest this example, and it is not clear that he would agree with this suggestion.

⁷ Deaton (2010) disputes the claim that randomized designs should be considered a ‘gold standard’ and argues that they do not deserve any special place in the hierarchy of approaches to estimate causal effects.

Whether existing estimates of intervention effectiveness provide useful evidence to guide future SDH interventions is a difficult question. In practice, the general approach is to combine evidence on effectiveness from multiple sources. Depending on the state of knowledge, this might involve expert judgment or systematic meta-analysis. Meta-analysis and structural econometric models are technically sophisticated approaches to generalize evidence on intervention effectiveness. However, for practical guidance to policy-makers considering alternative SDH interventions, another important criterion is that methods should be transparent. For the sake of transparency, it is useful to always present estimates of effectiveness and causal impacts in “natural units” as directly measured in the evaluation. The resulting lists of programme impacts will be intervention-specific, but have the advantage of transparency because they rely on as few assumptions as possible.

A further complication for evaluations of SDH interventions involves the time frame for measuring the causal impacts. Because of their potential importance, evaluations of SDH interventions should attempt to develop estimates of the long-term impacts. One possible approach parallels the methods used in some economic evaluations of clinical interventions, where a similar problem arises. Many RCTs of new pharmaceuticals measure short-term impacts in terms of clinical endpoints. For example, an RCT might estimate the impact of a new hypertension medication using the ‘natural units’ of measured blood pressure. However, to guide health policy questions such as whether public insurance should cover the new medication, an economic evaluation requires estimates of the impact on morbidity and mortality. In this situation, economic evaluations use ‘epidemiologic models’ from other studies to link the clinical endpoints to the policy-relevant endpoints. Magnuson suggests a similar two-step process to estimate longer-term impacts of early childhood interventions: “first link an early childhood outcome, such as achievement at age 5, to a more proximate outcome, such as adolescent achievement or high school graduation. The latter outcome could then be linked to an outcome of interest, such as adult earnings.” (National Academies 2009, p. 32).

The discussion above points to multiple sources of uncertainty in estimates of the effectiveness of SDH interventions: statistical uncertainty in experimental or quasi-experimental estimates of programme impacts; additional uncertainty if estimates from multiple sources are combined; and uncertainty in predictions about long-term impacts. Ideally, in a framework for economic evaluation of SDH interventions these sources of uncertainty should be accounted for and reflected in the results of the analysis. For example, after noting that previous estimates of the rate of return from the Perry Preschool Program are presented without standard errors, Heckman et al (2010) develop standard

errors for estimates of the rates of return and other impacts where meaningful standard errors can be determined. For the remaining components, Heckman et al carefully examine the sensitivity of estimates to plausible ranges of alternative assumptions. Similar strategies are used in economic evaluations of clinical interventions, for example to account for both the statistical uncertainty in the estimated treatment effect from an RCT, and the modelling uncertainty in linking clinical endpoints to policy-relevant endpoints (Gold et al 1996). To the extent possible, these strategies should be used to incorporate uncertainty into economic evaluations of SDH interventions.

3. Valuing the consequences of SDH Interventions

3. Valuing the consequences of SDH Interventions

Applied welfare economics provides a strong conceptual foundation for economic evaluations of SDH interventions. The common sense idea is that SDH interventions yield benefits because they make individuals in society happier. In economic terminology, SDH interventions increase individuals' utility (happiness or satisfaction), and social welfare is some aggregation of the utility levels of all individuals in the society. In cost-utility analysis, a form of cost-effectiveness analysis, health benefits are measured based on individual preferences over different health states, summarized in measures such as the quality-adjusted life year (QALY). In social cost-benefit analysis, social benefits are measured based on individuals' willingness to pay for the desirable outcome. Put differently, both methods try to value the health consequences: cost-effectiveness analysis uses a health metric while cost-benefit analysis uses a money metric.

3.1 Valuing costs

Both cost-effectiveness analysis and cost-benefit analysis require measuring the costs of the intervention under consideration. The concept of costs in economic evaluation is based on the same fundamental principle as the concept of benefits: SDH interventions create costs because they make some individuals in society unhappier. The link between SDH interventions and unhappiness, in turn, is the concept of opportunity cost. When resources are used in an SDH intervention, the resources can not be used in the production of other goods and services. Individuals are less happy because they have to give up the opportunity to consume these other desirable goods and services. Under conditions that often hold, the market prices of the resources used in an SDH will be a good measure of the opportunity costs. Standard references such as Boardman et al (2006) contain in-depth discussions of the challenges of measuring opportunity costs. Many of the challenges in measuring the costs of SDH interventions are similar to those discussed in the standard references.

In a recent review, Weatherly et al (2009) suggest that measuring intersectoral costs poses special methodological challenges for economic evaluation of public health interventions. Intersectoral costs pose the same challenges for economic evaluations of SDH interventions. The challenges stem from the fact that SDH interventions often have wide-ranging impacts so their costs may fall on individuals as well as on various parts of the public sector. Moreover, there may be ripple effects across different sectors. Weatherly et al (2009, p. 87) use an example which could be considered an SDH intervention:

“improvements in housing could reduce illness and injuries, with consequent reductions in health-care utilization.” The goal of a complete economic evaluation is to value all of the changes in resource use caused by the intervention. A complete evaluation should take into account whether a housing improvement reduces (or increases) health-care sector costs, while taking care to avoid double-counting costs or benefits.

Cost savings estimates have rhetorical appeal in discussions of a variety of public policies, including SDH interventions. The cost savings approach focuses on the impact of the intervention on either costs in the health care sector, or on public sector budgets. Cost savings are a component of the benefits of an SDH intervention. However, there are not conceptual grounds for focusing solely on this component and neglecting the other ways an SDH intervention improves social welfare. Indeed, a narrow focus on the health care sector or on public sector budgets can be quite misleading about the societal desirability of SDH interventions. Various SDH interventions that increase longevity might conceivably lead to higher lifetime health care costs, which in some countries will also translate into higher public sector costs. For example, precisely because of the heavy burden of diseases related to tobacco use, tobacco control efforts could actually increase future health care costs. Clearly, SDH interventions should not be judged failures because they are so successful in improving longevity that they increase lifetime costs. Instead, a complete economic evaluation (either cost-effectiveness or cost-benefit analysis) must be conducted to systematically compare all of the costs of the intervention to all of the benefits.

3.2 Cost-Effectiveness and Cost-Utility Analysis

Cost-effectiveness analysis (CEA) and cost-utility analysis (CUA) are widely used and accepted for economic evaluations of health interventions. CEA relates the costs of an intervention to a simple, common effect, often measured in natural units. For example, an evaluation of the protocol of guaiac tests for colon cancer estimated the costs per cancer detected. CUA is a form of CEA where the effect or outcome of health interventions are measured in a common metric based on people's utility levels or preferences over different health states. Probably the most popular common unit of measurement is the quality adjusted life year, or QALY, but there are other variants including the healthy year equivalent (HYE) and the disability adjusted life year (DALY). Using a common metric allows comparisons of a wide range of interventions.

CUA is a very well-established method for the economic evaluation of health care interventions. CUA relies on stated-preference methods to elicit preferences over different

health states. For example, in the standard gamble method, respondents are asked about their preferences between a gamble that might payoff in perfect health or might result in death, versus experiencing a sub-optimal health state (such as a chronic illness) with certainty. This and other methods are described in various standard references, including Gold et al (1996) and Drummond et al (1997). Because of its popularity, there are now many estimates of QALY weights that measure preferences over a wide range of health states.

For the purposes of this report, an important weakness is that CEA and CUA are hard to apply to the multiple impacts of SDH interventions.⁸ For example, a CEA or CUA of the health effects of an early childhood intervention would have to somehow incorporate the additional value of reductions in delinquency and crime due to the intervention. Some research is moving the QALY approach in this direction. Dolan et al (2005) extend the QALY approach to incorporate the intangible victim costs of violent crime. For example, they estimate that a murder results in about 18 QALYs lost, while a serious wounding results in 0.19 QALYs lost. Dolan and Peasgood (2007) further extend the approach to incorporate the costs of the fear of crime among potential victims.

Alternatively, the Institute of Medicine proposes a method to apply CUA to the analysis of regulations that yield both health and non-health impacts. This method calculates the comprehensive CUA ratio as the cost net of health care costs savings and other benefits per QALY saved. To net out the other benefits, this method requires willingness to pay estimates for all of the non-health impacts of the intervention. Put differently, this method requires a cost-benefit analysis for all non-health impacts, which is then integrated into the QALY-based CUA.

3.3 Cost-Benefit Analysis

Cost-benefit analysis (CBA) is based on societal willingness to pay for the health improvements and other consequences of SDH interventions. Like CEA, CBA relies on well-developed methods for estimating willingness to pay for health, especially mortality risks. One general approach is to use methods based on revealed preferences. Revealed preference methods analyze market behaviour to infer willingness to pay for non-market outcomes. For example, analysis of workers' choices about job safety and wages provide the basis for estimating the marginal value of mortality risks, often summarized as the "value

⁸ French et al (2008, p. 273) make a similar point about economic evaluations in addiction research, noting that "the variety and complexity of outcomes in addiction research...make it difficult to express economic impact through only one outcome, such as quality-adjusted life years (QALYs) gained."

of a statistical life.” More precisely, these studies estimate the dollar value people place on a small reduction in the risk of death. An example is useful to explain the terminology. Suppose an SDH intervention in the housing sector improves safety and reduces the risk of accidental death, say by 1 in 10 000. If each of 10 000 people are willing to pay US\$ 600 for that risk reduction, on aggregate the net benefits of the risk reduction are valued at US\$ 6 million. Because the intervention can be expected (in a statistical sense) to save one life, US\$ 6 million is called the statistical value of life. The same approach is also used to estimate willingness to pay to reduce the risks of on-the-job injuries.

The extensive research on the statistical value of life is reviewed and summarized in Viscusi and Aldy (2003). They review more than 60 studies that provide estimates of willingness to pay to reduce mortality risks, and about 40 studies that provide estimates of willingness to pay to reduce the risks of injuries. Estimates of willingness to pay to reduce mortality risks are available for at least 10 different countries. These estimates are directly relevant to the health benefits from various SDH interventions.

Instead of studying revealed preferences for health in labour and other markets, an alternative approach to estimate willingness to pay for health is to use stated preferences. Stated preference methods use contingent valuation surveys to directly elicit willingness to pay for non-market outcomes, including health. The methodology of contingent valuation surveys has been extensively studied and refined. Much of the research on the contingent valuation method concerns the application of the method to value environmental quality. In an important legitimization of the method, a “blue ribbon” panel of social scientists convened by the National Oceanic and Atmospheric Administration concluded that the contingent valuation method could provide useful estimates for the assessment of damages to natural resources (Arrow et al 1993). Standard references such as Boardman et al (2006, Chapter 14) provide in-depth discussions of the contingent valuation method and its strengths and weaknesses.

An important advantage of CBA compared to CUA is that it is straight-forward to apply to the multiple impacts of SDH interventions.⁹ The potential impacts of SDH interventions on social welfare include improvements in life expectancy, health-related quality of life, cognitive development, behaviour, and social competence, educational attainment, earnings, and reduced delinquency and crime. Economic methods have been developed to estimate willingness to pay for many of these outcomes.

⁹ Homer, Drummond and French (2008, p. 536) reach a similar conclusion about economic evaluation methods in addiction research. They argue that CBA is best suited to capture the societal benefits of substance abuse treatment, such as reduced criminal activity. Zavala et al (2005) provide a detailed discussion of CBA of adolescent substance abuse treatments, including illustrative estimates of the dollar value attached to outcomes related to education and employment, criminal activity, and juvenile justice services.

CBA of SDH interventions can be conducted using either a bottom-up or top-down approach to valuation. In the bottom-up approach, a dollar value is placed on each impact of the intervention, based on estimated willingness to pay for each outcome. The total benefits of an SDH intervention equal the total willingness to pay for all of the impacts. For example, in a bottom-up approach to place a monetary value on saving a high risk youth, Cohen and Piquero (2009) use estimates of three components of crime costs – victim costs, criminal justice system costs, and the lost productivity of offenders who are incarcerated. The value of these components is then added up to place a dollar value on preventing various criminal offenses. Cohen and Piquero then use these estimates to calculate the present value of the costs imposed by a career criminal, which forms the basis for their estimate of the value of saving a high risk youth. To continue the bottom-up approach, the value of saving a high risk youth might be one component of the benefits of an SDH intervention such as early childhood education. The value of preventing career criminals would then be combined with the value of the participants' higher earnings, improved health, and other outcomes.

The top-down approach to valuation uses estimates of willingness to pay for an impact at a higher level of aggregation. For example, Cohen and Piquero (2009) compare the bottom-up estimates of the value of the components of the costs of crime with top-down estimates of willingness to pay to prevent crime. To illustrate, in their bottom-up approach they estimate that each murder results in US\$ 4.6 million of victim costs, US\$ 300 000 in criminal justice system costs, and US\$ 140 000 in offender productivity losses, for a total of US\$5 million. In their top-down approach, Cohen and Piquero use an estimate that the willingness to pay to prevent a murder is US\$ 11.8 million.

In principle, a CBA of an SDH intervention could use the top-down approach based on direct estimates of willingness to pay for the intervention. Most applications of this approach would probably have to rely on stated preferences through a contingent valuation survey.¹⁰ For example, surveys could elicit willingness to pay for an early childhood intervention programme of a specified size. The method of conjoint analysis could enhance the usefulness of the survey. By presenting respondents with different scenarios, a conjoint analysis could provide estimates of willingness to pay for a range of programme sizes and other programme attributes. The obvious advantage of the top-down approach is that it is no longer necessary to piece together the values of all of the various component impacts of the SDH intervention.

¹⁰ In principle, it might be possible to use the revealed preference approach and infer willingness to pay for some SDH interventions based on market behaviours. For example, the value of early childhood education programs in a school system might be reflected in higher housing values. In this example, the challenge is to disentangle the value homeowners place on the programme for their own children, versus the value they place on the programme as an SDH intervention that helps disadvantaged children in their community. This might be possible by focusing on housing prices paid by childless homeowners.

It is also possible to take an intermediate position between bottom-up and top-down approaches to the valuation of SDH interventions. This approach could focus on proximate outcomes common to a variety of different SDH interventions, such as improvements in children's cognitive and non-cognitive abilities. Estimates of willingness to pay for improved child abilities could be useful to estimate the benefits of a range of SDH interventions, including early childhood education but also housing and environmental interventions.

3.4 Recommendations

The method of social cost-benefit analysis provides the most comprehensive approach to evaluate SDH interventions. Although cost-effectiveness analysis has gained widespread acceptance as a method to evaluate clinical interventions, its narrow focus on health effects makes it less suitable as method to evaluate SDH interventions.

Cost-benefit analysis methods require the use of time discounting and methods to account for uncertainty (see, for example, Boardman et al 2006, Chapters 6, 7, 8 and 10). These are mainly standard issues that do not pose special conceptual challenges for the economic evaluation of SDH interventions. Time discounting, however, is an important practical consideration. Many potential SDH interventions, such as early childhood interventions, are investments that involve immediate costs that only yield a stream of benefits 10, 20, or more years in the future. In this context, the choice of a discount rate is a very important determinant of whether the discounted present value of the net benefits of the intervention.

4. Valuing reductions in health inequities

4. Valuing reductions in health inequities

Both the theory and practice of economic evaluation tend to shift the focus away from the value of reductions in health inequities. The theoretical foundations of economic evaluation methods focus on efficiency – the total health gains or the total benefits from interventions, regardless of how they are distributed across different members of society. In practice, recent literature reviews have confirmed that to date economic evaluation studies in health care (Sassi, Archard and Le Grand, 2001) and in public health (Drummond et al 2006) have for the vast majority chosen not to take into account distributional effects. At the same time, several empirical studies have shown that people are willing to sacrifice overall health benefits for a reduction in health inequalities (Dolan and Shaw 2001).

The emphasis on efficiency over equity tends to neglect one of the central goals of SDH interventions. This section reviews approaches to incorporate the value of reductions in health inequities into economic evaluations of SDH interventions. To set the stage, we begin with a discussion of equity-efficiency trade-offs, before turning to more detailed discussions of the role of health inequities in cost-effectiveness and cost-benefit analysis.

4.1 The equity-efficiency trade-off

A long tradition in welfare economics makes the conceptual distinction between the policy goals of improving equity versus efficiency. Economists also stress that in many contexts policy-makers face a trade-off between equity and efficiency. For example, a progressive system of taxation and income transfers that leads to a more equal distribution of income and wealth also creates work dis-incentives and economic efficiency costs. The economic evaluation method of social cost-benefit analysis evaluates interventions' impacts on economic efficiency, not equity. But by quantifying the effects on economic efficiency, cost-effectiveness analysis and cost-benefit analysis can quantify the equity-efficiency trade-offs in different policy alternatives. This section discusses when such trade-offs might be desirable.

At the outset, however, it should be noted that such trade-offs are not inevitable: some SDH interventions address both equity and efficiency. In fact, Nobel Laureate James Heckman argues forcefully that: "Investing in disadvantaged young children is a rare public policy with no equity-efficiency trade-off. It reduces the inequality associated with the accident of birth and at the same time raises the productivity of society at large." Heckman and Masterov (2007) make the case by reviewing "substantial evidence ... that these children are more likely to commit crime, have out-of-wedlock births and drop out of school.

Early interventions that partially remediate the effects of adverse environments can reverse some of the harm of disadvantage and have a high economic return. They benefit not only the children themselves, but also their children, as well as society at large.” In this situation, the emphasis on efficiency in economic evaluation goes hand in hand with the policy objective of reducing inequities.

When policy makers face an equity-efficiency trade-off, the results of cost-effectiveness analysis or cost-benefit analysis provide guidance by quantifying the efficiency losses incurred to improve equity. Equity-efficiency trade-offs are probably still common among SDH interventions. While Heckman and Jacobs (2010) continue to stress that investments in early childhood do not face an equity-efficiency trade-off, they argue that: “there is a substantial trade-off at later ages. Later remediation of skill deficits acquired in early years often does not meet the cost-benefit criterion.” In terms of health inequities, the US experience with the Healthy People 2010 public health initiative is instructive. Healthy People 2010 sets out two major goals: to increase quality and years of healthy life; and to eliminate health disparities. However, Mechanic (2002) and Keppel et al. (2007) stress that progress towards these two goals may not necessarily coincide. Based on the mid-course review of the Healthy People 2010 initiative, Keppel et al. report that, for 69 specific objectives, the outcome was progress towards increasing the quality and length of life with little or no change in relative disparity. For ten objectives progress towards the two goals actually moved in opposite directions.

Quantifying the equity-efficiency trade-off is useful to the extent policy-makers are willing to make such trade-offs. The relative importance of equity versus efficiency involves deep ethical and philosophical issues. But at a practical level, most observers agree that neither goal is of such paramount importance that all trade-offs are undesirable. To take a hypothetical example of a cost-effectiveness analysis, suppose a policy-maker has to choose between two interventions A and B, which have the same costs and benefit the same number of people. Intervention A improves the life expectancy of a group of high-income individuals by 10 years, while intervention B improves the life expectancy of a group of disadvantaged low-income workers by 9 years. Although intervention A is more cost-effective, many policy makers would probably agree that intervention B is the better policy choice. Similarly, discussions of cost-benefit analysis typically include the caveat that interventions that fail the cost-benefit criterion of efficiency might still be desirable because of their implications for equity. For example, suppose a worker training programme in later life does not pass the cost-benefit criterion because it imposes US\$10 million of costs on upper-income taxpayers but only yields US\$9 million of benefits to the disadvantaged low-income

trainees. Many policy-makers would probably agree that this efficiency cost is worth incurring to achieve a more equitable outcome.

By the same token, most policy-makers would probably agree that policies that lead to greater inequities can still be socially desirable. Deaton (2002) provides a strong argument for this view: “Consider a technical innovation - for example, a new life-saving procedure or new health-related knowledge. Coronary artery bypass grafts or neonatal intensive care units are good examples of the former; for the latter, think of the surgeon general’s report on smoking in 1964 or the application at the turn of the twentieth century of the germ theory of disease to personal and medical hygiene. Better-educated people will be quicker to adopt or benefit from the innovation; if the innovation is not immediately available to everyone, money might help, too. Because the innovations are beneficial to health, some people’s health is improved and other people’s health stays the same or is improved less. Because of the role of education and income, the gradient steepens; the health of the rich and well educated improves more. The Pareto [cost-benefit] criterion says that such innovations are beneficial and are to be encouraged.” Deaton rejects the arguments that all “inequalities are inherently bad, and innovations that increase them are to be discouraged. Policies based on such arguments are misconceived; they result in some people dying who could have lived, without preventing any other deaths.”

Williams and Cookson (2000) provide an in-depth discussion of the philosophical notions of justice behind the concern about equity and health. They also review attempts to quantify the tradeoffs between health equity and efficiency, although they conclude that the empirical work “is still at an embryonic stage...no-one has yet devised a fully satisfactory questionnaire method for eliciting equity-efficiency tradeoffs in health.” They go on to stress that this lack of reliable data applies more generally to equity-efficiency trade-offs in other sectors.

4.2 Valuing reductions in health inequities in Cost-Effectiveness

Analysis

QALY-based cost-effectiveness analysis (also known as cost-utility analysis) is often justified as a tool to help decision-makers maximize the health gains possible given a constrained health sector budget. Any decision made solely on the basis of maximizing health gains across the whole population is one which implicitly gives an equal weight to one

QALY gained regardless of who gains it.¹¹ This “a QALY is a QALY” assumption can only be valid for decision-making if society has no desire to give additional weight for QALYs that are accrued to certain groups of patients. These could be the socioeconomically deprived, the young, those who are severely ill, or those who have a very low life expectancy. For example, the “a QALY is a QALY” assumption rules out the possibility that society might prefer to deliver an improvement of 0.1 in the quality of life state to a patient who is severely ill over someone who is in near perfect health (that is a preference for an improvement in the quality of health state of one patient from 0.2 to 0.3 rather than an improvement in another patient from 0.8 to 0.9). Any divergence away from the strict “a QALY is a QALY” assumption, in order to value more equitable outcomes, will be associated with some loss in total health outcomes. This is an example of the equity-efficiency trade-off.

Although it neglects health inequities, it should be noted that the “a QALY is a QALY” assumption helps rule out decisions that lead to outcomes which directly contradict fundamental ethical, legal and political principles. This is particularly true for clinical decisions, where the health economic evaluation literature is currently concentrated, but does not necessarily hold for public health interventions. It is hard to envisage any situation where it would be acceptable to actively deny a clinical intervention, such as a cholesterol-lowering statin therapy, to a particular group of patients based on their income or place of residence. Yet this kind of discrimination, in the form of targeting certain groups, is at least acceptable, and may even be actively encouraged, in public health policy.

Equity-weighting analysis has been proposed as an extension of cost-effectiveness analysis that incorporates the value society places on reductions in health inequities (Williams and Cookson 2006). The basic idea of the approach is to reject the “a QALY is a QALY” assumption and explicitly place greater weight on the QALY gains of certain groups. In this approach, it would be possible to place a weight on the QALY gains associated with competing public health policy according to the effect that it has on health inequities. For example, a QALY gain that is also associated with a reduction in health inequities would be given more weight than a QALY gain than a QALY gain that leaves health inequities in place. Ultimately, alternative policies could then be ranked, rather neatly, not just in terms of their cost per QALY, but in terms of some cost per ‘equitable-QALY’ measurement.

While the need to address equity concerns in economic evaluation is widely shared, there are a number of unresolved issues with the ‘equity weighting’ approach (Richardson

¹¹ Also regardless of how the QALY is achieved. Society may also have other concerns, in addition to equity considerations, such as a desire to protect individual liberty in terms of the freedom to make choices. In the United Kingdom, for example, a recent meeting of a citizen’s council indicated that the public consider that “non-mandatory public health measures, such as providing education and information, were preferable to mandatory ones, provided they were effective”. NICE Social Value Judgements (second edition, page 26.)

2009). Unlike monetary outcomes, equity outcomes could have a variety of conflicting meanings. It has been argued that there are so many potential dimensions of equity that the use of equity weights would substantially complicate the analysis (Tsuchiya, McCabe and Wailoo 2007). It seems unlikely that it would be sufficient to incorporate the equity weights at an aggregate level. More likely weights would need to be applied separately for each individual targeted by a policy according to a range of equity-relevant characteristics. In any case there is very little evidence of the approach having been used in practice.¹²

4.3 Valuing reductions in health inequities in Cost-Benefit Analysis

In standard social cost-benefit analysis, the net benefits of an intervention are calculated without regard to how the benefits and costs are distributed to different members of society. As a result, even though they generate large benefits for disadvantaged populations, reductions in health inequities will not necessarily lead to positive net benefits in a cost-benefit analysis. As explained by Arnold Harberger, one of the pioneers of social cost-benefit analysis, the equal weighting of benefits and costs without regard to their distribution is fundamentally “a technical convention which permits us to separate resource-allocation from distributional effects in the analysis of any given problem.” (Harberger 1974, p.3). Harberger further states that “I emphatically do not mean to say or imply thereby that distributional considerations are unimportant, or that economists should refrain from expressing opinions concerning them. In fact, I believe that such opinions can play a vital role in the public debate over many policy issues, especially on the wide range of programs with explicit distributional orientation.” SDH interventions are perfect examples of such programs.

Several approaches, including the use of distributional weights and the basic needs approach, have been proposed and implemented as ways to incorporate societal concerns about inequities. Analogous to equity-weights in cost-effectiveness analysis, the basic idea of distributional weights is to reject the “a dollar is a dollar” assumption made in standard cost-benefit analysis. Appropriate choice of distributional weights could reflect a willingness of society to redistribute income from the rich to the poor. Alternatively, the basic needs approach retains the assumption that “a dollar is a dollar,” but places higher dollar values on changes in the consumption of goods and services that meet basic human needs, such as health care, food, and shelter. Again, SDH interventions are perfect examples of programs that would be more highly valued using the basic needs approach to cost-benefit analysis.

¹² One notable exception, which includes explicit time and age weights, is: Murray CJL and Lopez AD. *The Global Burden of Disease*. Geneva, World Health Organization, 1996.

Harberger (1974) expresses doubts that economics could achieve a consensus about the weight that should be attached to the welfare of different groups. However, currently existing methods used for valuing different health states, such as discrete choice experiments, could be used to elicit information from the general public about their preferences for various equity-efficiency tradeoffs. For example, Cai, Cameron and Gerdes (forthcoming) use a stated preference survey to explore equity trade-offs in choices over policies to prevent climate change. They find that some respondents' willingness to pay is higher when they believe that the impacts of climate change may be borne disproportionately by the world's poor. Stated-preference surveys could similarly elicit willingness to pay for the reductions in health inequities due to SDH interventions. It would then be a straight-forward exercise to incorporate estimates of willingness to pay to reduce health inequities in cost-benefit analyses of SDH interventions.

4.4 Recommendations

Table 1 presents a hierarchy of approaches to incorporating equity considerations into economic evaluations of SDH interventions. The hierarchy is an adapted version drawing on Cookson, Drummond and Weatherly (2008). The first two approaches are recommended as relatively unproblematic approaches. Neither attempts an explicit valuation of society's willingness to pay for reductions in health inequalities but, if published alongside the results of standard cost-effectiveness or cost-benefit analyses, they could provide valuable supplementary information for decision makers who are concerned about the impact of competing health interventions on health inequities.

The first of these approaches is the simple exercise of bringing together all existing information relating to how each of the policy interventions under consideration might affect health inequities. Ideally this would include information about existing trends in a particular health inequity, how those health inequities have been affected by similar interventions elsewhere, and anything that is known about society's willingness to pay for a reduction in that health inequity. Clearly this is only a first step to incorporating equity considerations into economic evaluation, and in reality may disclose little relevant information. But, in terms of best practice, it should be considered a minimum requirement.

The second approach, health inequity impact assessment, is an attempt to quantify the impact that competing interventions are likely to have on various health inequities. This is likely to take the form of collecting data on how the cost-effectiveness or net benefits of interventions will change if the intervention is targeted at different sub-groups of the

population. Standard evaluation methods would be suitable to achieve this. However, clinical trials tend to measure average effects and lack the necessary data relating to specific sub-groups of people. So this technique is likely to be costly in terms of additional data requirements.

The choice of approach should be determined by a number of factors. Progression through the hierarchy is reliant on the availability of suitable information and data such that in many cases it may only be possible to complete the simple review of background information. However, it is also important that a decision to invest resources in completing one of the more advanced approaches should only be made with consideration of how valuable it is to the decision maker to incorporate a particular equity consideration into the economic analysis. Or, to put it another way, how sensitive is a decision that was made on the basis of QALY maximization or net benefit maximization methods to the inclusion of equity considerations?

Table 1: A hierarchy of approaches to incorporating equity considerations into economic evaluations of SDH interventions

Approaches	Advantages	Disadvantages
Review of background information on health inequities	<p>Cheap, easy and quick to complete.</p> <p>Ensures that all existing relevant materials are available to the decision maker.</p> <p>A useful first step.</p>	<p>No new methodology is required.</p> <p>Can only be used as an insight to some associated issues rather than providing a conclusive answer.</p>
Health inequity impact assessment	<p>Does not require any new methodology in terms of completing cost-effectiveness or cost-benefit analyses.</p> <p>Does not require measurement of how much society values changes in health inequities.</p> <p>Provides decision-makers with evidence about the cost of reducing health inequities.</p>	<p>Requires data on the cost-effectiveness or net benefits of an intervention by sub-group of the population, which is not often collected.</p>
<p>Equity weighting of health outcomes</p> <p>OR</p> <p>Willingness to pay for health inequities</p>	<p>The only approach which incorporates a quantifiable value for society's willingness to pay for a reduction in health inequalities.</p> <p>If completed at the individual level it seems likely that the approach could provide a very accurate tool for decision-makers.</p>	<p>Costly, in terms of time, data collection, and computing power.</p> <p>Not yet used in practice.</p> <p>A lack of suitable individual-level data.</p>

References

References

Angrist JD and Lavy V. Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics*, 1999, 114(2):533–75.

Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 1996, (434):444–455.

Angrist J and Pischke J-S. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton and Oxford, Princeton University Press, 2009.

Angrist J and Pischke J-S. The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *Journal of Economic Perspectives*, 2010, 24(2):3–30.

Arrow KJ, Solow R, Portney P, Leamer E, Radner R and Schuman. *Report of the NOAA Panel on Contingent Valuation*. Federal Register 58 (January 15, 1993):4601–4614.

Boardman AE, Greenberg DH, Vining AR and Weimer DL. *Cost-Benefit Analysis: Concepts and Practice*, 3rd ed. Prentice Hall, 2006.

Beile C, Cameron TA and Gerdes GR. *Distributional Preferences and the Incidence of Costs and Benefits in Climate Change Policy*, Environmental and Resource Economics [forthcoming].

Card D. Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, 2001, 69:127–160.

Carlson D, Haveman R, Kaplan T and Wolfe B. The benefits and costs of the Section 8 housing subsidy program: A framework and estimates of first-year effects. *Journal of Policy Analysis and Management*, Spring 2011, 30 (2):233–255.

Chetty R. Sufficient Statistics for Welfare Analysis: A Bridge between Structural and Reduced-form Methods. *Annual Review of Economics*, 2009, 1:451–488.

Cohen MA and Piquero AR. New Evidence on the Monetary Value of Saving a High Risk Youth. *Journal of Quantitative Criminology*, 2009, 25:25–49.

Cook TD and Campbell DT. *Quasi-Experimentation: Design and Analysis for Field Settings*. Chicago, Rand McNally, 1979.

Commission on Social Determinants of Health. *Closing the gap in a generation. Health equity through action on the social determinants of health*. Geneva, World Health Organization, 2008.

Cunha F and Heckman JJ. *The Economics and Psychology of Inequality and Human Development*. Working Paper 14695, National Bureau of Economic Research, 2009.

Deaton AS. The Policy Implications of the Gradient of Health and Wealth. *Health Affairs*, 2002, 21(2):13–30.

Deaton AS. Instruments, Randomization, and Learning about Development. *Journal of Economic Literature*, 2010, 48(2):424–455.

Deven C, Haveman R, Kaplan T and Wolfe B. The Benefits and Costs of the Section 8 Housing Subsidy Program: A Framework and First-Year Estimates. University of Wisconsin, *La Follette School Working Paper* No. 2009-025, 2009.

Dolan P, Loomes G, Peasgood T and Tsuchiya A. Estimating the Intangible Victim Costs of Violent Crime. *British Journal of Criminology*, 2005, 45:958–976.

Dolan P and Peasgood T. Estimating the Economic and Social Costs of the Fear of Crime.” *British Journal of Criminology*, 2007, 47:121–132.

Dolan P and Shaw R. *A review of people’s preferences regarding the equity-efficiency trade-off in health*. 2001, [Unpublished manuscript].

Drummond M et al. Assessing the challenges of applying standard methods of economic evaluation to public health interventions, Report prepared for the Department of Health Policy Research Programme, 2006.

Drummond M, O’Brien BJ, Stoddart GL and Torrance G. *Methods for the Economic Evaluation of Health Programmes*. Oxford, Oxford University Press, 1997.

Epstein D, Jimenez Rubio D, Smith PC, Suhrcke M. Social determinants of health: an economic perspective. *Health Economics*, 2009, 18:495–502.

Fair Society, Healthy Lives: The Marmot Review. Strategic Review of Health Inequalities in England post 2010. The Marmot Review, 2010.

French MT, Zavala SK, McCollister KE, Waldron HB, Turner CW and Ozechowski TJ. Cost-effectiveness Analysis of Four Interventions for Adolescents with a Substance Use Disorder. *Journal of Substance Abuse Treatment*, 2008, 34:272(1996)281.

Gold MR, Siegel JE, Russell LB and Weinstein M, eds. *Cost-Effectiveness in Health and Medicine*. New York, Oxford University Press, 1996.

Harberger AC. *Taxation and Welfare*. Chicago and London, The University of Chicago Press, 1974.

Heckman JJ. Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy. *Journal of Economic Literature*, 2010, 48(2):356–398.

Heckman JJ and Jacobs B. *Policies to Create and Destroy Human Capital in Europe*. NBER Working Paper 15742, 2010.

Heckman JJ and Masterov DV. The Productivity Argument for Investing in Young Children. *Review of Agricultural Economics*, 2007, 29(3):446–493.

Heckman JJ and Vytlacil E. Structural Equations, Treatment Effects and Econometric Policy Evaluation. *Econometrica*, 2005, 73(3):669–738.

Heckman JJ, Malofeeva L, Pinto R and Savelyev PA. *The Effect of the Perry Preschool Program on the Cognitive and Non-Cognitive Skills of its Participants*. [Unpublished manuscript] University of Chicago, Department of Economics, 2008.

Heckman JJ, Moon SH, Pinto R, Savelyev PA and Yavitz A. “The rate of return to the HighScope Perry Preschool Program.” *Journal of Public Economics*, 2010, 94:114–128.

Homer JF, Drummond MF and French MT. Economic Evaluation of Adolescent Addiction Programs: Methodologic Challenges and Recommendations. *Journal of Adolescent Health*, 2008, 43: 529–539.

Ichino A, Winter-Ebmer R. Lower and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments. *European Economic Review*, 1999, 43(4-6):889–901.

Imbens GW. Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature*, 2010, 48(2):399–423.

Imberman S, Kugler AD and Sacerdote B. *Katrina's children: evidence on the structure of peer effects from hurricane evacuees* (Working Paper No. 15291), National Bureau of Economic Research, 2009.

Institute of Medicine. *Valuing Health for Regulatory Cost-Effectiveness Analysis*. Washington DC, National Academies Press, 2006.

Jones A. Evaluating public health interventions with non-experimental data analysis, *HEDG, CHE Research Conference, May 2006*.

Karoly LA, Kilburn MR, Bigelow JH, Caulkins JP and Cannon JS. Assessing Costs and Benefits of Early Childhood Intervention Programs. RAND Corporation, Santa Monica, California, 2001.

Keppel K, Bilheimer L and Gurley L. Improving Population Health and Reducing Health Care Disparities. *Health Affairs*, 2007, 26:1281–1292.

Lavin T, Metcalfe O. Economic arguments for addressing social determinants of health inequalities. DETERMINE Working document #4, 2009.

Lee DS and Lemieux T. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 2010, 48(2):281–355.

Macintyre S. Evidence based policy making. *BMJ*, 2003, 326:5–6.

Mechanic D. Disadvantage, Inequality, and Social Policy. *Health Affairs*, 2002, 21(2): 48–59.

Meyer BD. Natural and Quasi-Experiments in Economics. *Journal of Business & Economic Statistics*, 1995, 13(2):151–61.

Mazzucco S, Meggiolaro S, Suhrcke M. *The economic benefits of reducing health inequalities in England and Wales*. Background report to the Marmot Review, 2010.

Murray CJL and Lopez AD. *The Global Burden of Disease. Vol.1*. Geneva, World Health Organization, 1996.

National Academies. Strengthening Benefit-Cost Analysis for Early Childhood Interventions: Workshop Summary, 2009.

Neumann PJ, Jacobson PD and Palmer JA. Measuring the Value of Public Health Systems: The Disconnect Between Health Economists and Public Health Practitioners. *American Journal of Public Health*, 2008, 98(12):2173–2180.

Nevo A. *Taking the Dogma Out of Econometrics: Structural Modeling and Credible Inference*. The Center for the Study of Industrial Organization at Northwestern University. Working Paper # 0104, 2010.

Nores M and Barnett WS. Benefits of early childhood interventions across the world: (Under) Investing in the very young. *Economics of Education Review*, 2010, 29(2):271–282.

Milagros N and Barnett WS. Benefits of Early Childhood Interventions Across the World: (Under) Investing in the Very Young. *Economics of Education Review*, 2010, 29:271–282.

Petticrew M, Cummins S, Ferrell C, Findley A, Higgins C, Hoy C, Kearns A, Sparks L. Natural experiments: an underused tool for public health? *Public Health*, 2005, 119:751–757.

Pillas D, Suhrcke M. *Assessing the potential or actual impact on health and health inequalities of policies aiming to improve Early Child Development (ECD) in England*. Background report to the Marmot Review, 2009.

Rosenzweig M and Wolpin K. Natural ‘Natural Experiments’ in Economics. *Journal of Economic Literature*, 2000, 38:827–874.

Richardson J. *Is the incorporation of equity considerations into economic evaluation really so simple? A comment on Cookson, Drummons and Weatherly*, Health Economics, Policy and Law, 4, p.247-254, 2009.

Saelens BE, Sallis JF and Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Annals of Behavioral Medicine*, Spring, 2003, 25(2):80–91.

Sassi FL, Archard L and Le Grand J. Equity and the economic evaluation of healthcare. *Health Technology Assessment*, 2001, 5(3):1–130.

Social Value Judgements: Principles for the Development of NICE Guidance, 2nd ed., London, National Institute for Health and Clinical Excellence.

Tsuchiya A, McCabe C, Wailoo A. *Incorporating equity weights into cost-effectiveness analysis: opening Pandora’s box?* Paper written for the Health Economics Study Group meeting, University of Birmingham, January 2007.

Viscusi W and Aldy J. The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World. *Journal of Risk and Uncertainty*, 2003, 27(1): 5–76.

Wanless D. *Securing good health for the whole population: final report*. London, HM Treasury, 2004.

Weatherly et al. Methods for assessing the cost-effectiveness of public health interventions: Key challenges and recommendations. *Health Policy*, 2009, 93(2):85–92.

Wilensky GR and Satcher D. Don’t Forget About the Social Determinants of Health. Health Affairs Web Exclusive, 16 January 2009, w194 - w198.

Williams AH and Cookson RA. Chapter 35 Equity in Health. In: Culyer AJ and Newhouse JP, eds. *Handbook of Health Economics*. Elsevier, Volume 1, Part B, 2000:1863–1910.

Williams AH and Cookson RA. Equity-efficiency trade-offs in health technology assessment. *International Journal of Technology Assessment in Health Care*, 2006, 22(1):1–9.

Zavala S, French MT, Henderson CE, Alberga L, Rowe C and Liddle H. Guidelines and Challenges for Estimating the Economic Costs and Benefits of Adolescent Substance Abuse Treatments. *Journal of Substance Abuse Treatment*, 2005, 29:191–205.

The WHO Regional Office for Europe

The World Health Organization (WHO) is a specialized agency of the United Nations created in 1948 with the primary responsibility for international health matters and public health. The WHO Regional Office for Europe is one of six regional offices throughout the world, each with its own programme geared to the particular health conditions of the countries it serves.

Member States

Albania
Andorra
Armenia
Austria
Azerbaijan
Belarus
Belgium
Bosnia and Herzegovina
Bulgaria
Croatia
Cyprus
Czech Republic
Denmark
Estonia
Finland
France
Georgia
Germany
Greece
Hungary
Iceland
Ireland
Israel
Italy
Kazakhstan
Kyrgyzstan
Latvia
Lithuania
Luxembourg
Malta
Monaco
Montenegro
Netherlands
Norway
Poland
Portugal
Republic of Moldova
Romania
Russian Federation
San Marino
Serbia
Slovakia
Slovenia
Spain
Sweden
Switzerland
Tajikistan
The former Yugoslav Republic of Macedonia
Turkey
Turkmenistan
Ukraine
United Kingdom
Uzbekistan

WHOLIS number
Original: English

World Health Organization Regional Office for Europe

Scherfigsvej 8, DK-2100 Copenhagen Ø, Denmark
Tel.: +45 39 17 17 17. Fax: +45 39 17 18 18. E-mail:
contact@euro.who.int
Web site: www.euro.who.int