C Inter-generational comparisons: the social welfare function approach

C.1 Inter-generational comparisons: the representative consumer

Discounting costs and benefits received in the far future involves valuing the effects of policies on future generations, raising ethical issues. For example, the current generation can adopt policies which harm future generations – not only a different group of people, but one that is not around to defend its interests.

Discounting in the long-term involves comparing the welfare of different generations. The usual approach is to do so explicitly with a representative, infinitely–lived household who maximises the following utility function:[1]

$$U = u(c_0) + (1+p)^{1-\varepsilon}\gamma u(c_1) + (1+p)^{2(1-\varepsilon)}\gamma^2 u(c_2) + ... = \sum_{j=0}^{\infty}(1+p)^{j(1-\varepsilon)}\gamma^j u(c_j)$$

where c_t is the per person consumption of a typical household member during period t. Each period represents a different generation. The population (and so representative household) grows at the exogenous rate *p*. $\gamma = \frac{1}{(1+\theta)}$ with $\theta > 0$ is the pure rate of time preference, which is used to discount the utility of future generations..(1-ε) is the weighting the representative gives to the number of people in each generation.

There are two possible ways to proceed using this function: a prescriptive (normative) approach or a descriptive (positive) approach. The prescriptive approach interprets the utility function of the infinitely lived household as an inter-generational social welfare function. The social welfare function represents some ethical judgement about the appropriate distribution of welfare across generations. Just as a utility function shows how a person ranks different combinations of consumption goods, the social welfare function represents a value judgement of

1 See, for example, McCallum (1996).

how society should rank different distributions of utility across generations. It formalises the trade–off between efficiency and equity. A social welfare function gives the welfare of the whole society as a function of the utilities of individuals, just as the utility function gives the welfare of the individual as a function of the quantities of goods the individual consumes.

The descriptive approach uses this function as positive model of how the economy actually works and how the interest rate is determined. For example, a competitive economy with overlapping generations could operate in this manner (with c_t being the value of lifetime consumption of generation t).

Special cases of the utility function are often used under both approaches. For example, if ε=1, then $U = \sum_{j=0}^{\infty} \gamma^j u(c_j)$. Utility depends only on the average consumption of family members and the number of descendants does not affect the representative consumer's utility.[2] Alternatively, a constant population (p=0) would give the same result, and now models a constant population of identical individuals with the same utility function.

If ε=0, then $U = \sum_{j=0}^{\infty} (1+p)^j \gamma^j u(c_j)$. This has been called a 'Benthamite' formulation, where the household maximises the present value of the total utility of all current and future household members. Because the household cares about both the average utility and number of descendants, when the population is growing, it cares more about the future than a household that cares only about its descendants' average utility (an increase in future consumption per person has a larger effect on utility).

C.2 Deriving the discount rate

The production side of the economy is a standard Solow model, with exogenous labour–augmenting (Harrod neutral) technical progress at rate g. Output per household each period is a constant returns function of per household capital K and efficiency units of labour, N: $Y_j = Y(K_j, N_j)$ where $N_j = (1+g)^j (1+\pi)^j$. That is, exogenous labour–augmenting (Harrod neutral) technical progress increases the number of efficiency units of labour per head of population at rate g.

Now express the production function in per efficiency unit terms: let $y_j = Y_j/N_j$ and $k_j = K_j/N_j$ and $y_j = (1/N)Y(K/N,1) = y(k)$.

2 For example, Blanchard and Fischer (1989, pp. 38-39) use this case.

In the steady state, the effective labour supply, real income and the capital stock all grow at rate $g + p + gp$. Therefore Y/K is constant and so is the rate of return on capital, $Y_K = \partial F/\partial K = \partial f/\partial k$.[3] The share of income going to capital, $Y_K K/Y$ is constant. The share going to labour is constant. The distribution of income is steady, which requires Harrod neutral technical progress.[4]

In the steady state, output, consumption, the wage and capital per efficiency unit are constant. But it is people, not efficiency units, that receive income and consume. Population is growing at rate p. Consumption, capital, the wage and output per person grow at rate g, the rate of technical progress.

A steady state with c_t growing at rate g, is only possible if u′ has a constant elasticity with respect to c_j.[5] Assume, therefore, that the household has a power utility function: $u(c_j)=\dfrac{c_j^{1-\eta}}{1-\eta}$ with $\eta > 0$, which gives $u(c_j) = ln\ (c_j)$ for $\eta = 1$. Power utility has a constant elasticity of marginal utility of $-\eta$ (which is also the co-efficient of relative risk aversion). It determines the rate at which an individual's marginal utility falls as income rises.

The household's budget constraint is:

$$y(k_j) = c_j + (1+p)k_{j+1} - (1-d)k_j$$

where capital depreciates at rate d.

The household chooses values of c_j and $k_1, k_2, \ldots$ to maximise utility:

$$u(c_0) + (1+p)^{1-\varepsilon}\beta u(c_1) + (1+p)^{2(1-\varepsilon)}\beta^2 u(c_2) + \ldots$$

subject to the budget constraint and given k_0.

The first order conditions give: $(1+\theta)(1+p)^{\varepsilon}(c_j/c_{j+1})^{-\eta} = y'(k_{j+1}) + 1 - d$

That is the interest rate is:

$$r = y'(k) - d = (1+\theta)(1+p)^{\varepsilon}(1+g)^{\eta} - 1$$

Or approximately (true when expressed as continuously compounded rates):

$$r = y'(k) - d = \theta + \varepsilon p + \eta g$$

3 $\partial F/\partial K = \partial(Nf)/\partial K = N(\partial k/\partial K)(\partial f/\partial k) = \partial f/\partial k$

4 In the case of a Cobb-Douglas production function, Harrod neutral technical progress is also Hicks neutral. $Y = K^{\alpha}[a(t)L]^{1-\alpha} = [a(t)]^{1-\alpha}K^{\alpha}L^{1-\alpha}$

5 That is, $c_j u''(c_j)/u'(c_j)$ is constant. see McCallam (1996, footnote 11, p. 46).

In equilibrium the net return on capital equals the marginal rate of time preference equals the interest rate. As there are no taxes, the consumer equals the investment rate. If the prescriptive approach specifies a Ramsey rate that is below the marginal rate of return on investment, welfare would be improved by increasing the capital stock until the equilibrium condition holds.

We have derived the Ramsey formula for the social discount rate accounting for population growth: $r = \theta + \varepsilon p + \eta g$ where θ is the pure rate of time preference used to discount utility. ε depends on how changes in the population are valued and determines the effect of population growth, p on the discount rate. Consumption per person grows at rate g and η is the (absolute value of) the elasticity of the marginal utility of consumption (the percentage fall in the marginal utility when consumption increases by one per cent). Most authors ignore the effect of population growth on the discount rate — assuming either constant population (p=0) or a Benthamite social welfare function (ε=0), which gives a lower discount rate (when population growth is positive).

The interest rate r is the appropriate rate to discount consumption. The consumption discount rate is higher than the utility discount rate because the growth in per head consumption means the marginal utility of future consumption is less than current consumption (and η shows the rate at which it falls). A unit of consumption now gives more utility than a unit of consumption in the future.

C.3 The prescriptive approach

The prescriptive approach uses the Ramsey formula as a framework to guide the ethical choice of a social discount rate. It involves specifying the parameters to reflect some ethical beliefs, sometimes literally appealing to philosophers. There is much disagreement about these parameters. The prescriptive approach requires a subjective assessment of them and risks the analyst imposing his own judgement.

The parameter θ is now the pure social rate of time preference,[6] which determines the relative weights put on the welfare of different generations in the social welfare function, so so-called distributional weights approach. A positive parameter means less weight is put on the utility of future generations. One reason for a positive θ is the chance of some catastrophic event eliminating human life on earth. More consumption for future generations is worth less if there is some chance they will not exist. The probability of extinction is usually considered quite low. For

6 Some authors call this the social rate of time preference, but that is also used to refer to the social discount rate (which is used to discount consumption, not utility).

example, Stern sets it at 0.1 per cent (that is an annual probability of one in a thousand).

The other reason for a positive θ is if utility to future generations counts less than utility to the current generation. Some authors argue that θ should be set to zero (that is, $\gamma = 1$) so that all generations count equally. For example, Stern argues that it is ethical to give the same weight to the utility of different generations. Individuals often discount their own future utility because of impatience. Cowen argues that utility is not 'productive' over time as is invested capital and that impatience is not relevant in an inter-generational setting because future generations are not impatient to be born — and they do not experience a disutility of waiting to be born.

Others disagree and argue that if future generations are better off than current generations, their utility should be discounted. The appropriate weight to put on the interests of different generations involves comparing the welfare of different people (inter-personal justice). There is a well–developed literature on use of social welfare functions to compare people's welfare. As Brennan (2005) points out, Stern's preferred social welfare function has the Benthamite form:

$$U = u(c_0) + (1+\pi)u(c_1) + (1+\pi)^2 u(c_2) + ... = \sum_{j=0}^{\infty} (1+\pi)^j u(c_j)$$

where social welfare is simply the sum of the individual utilities.

This additive social welfare function, which maximises the sum of the utilities is often described as utilitarian, representing Bentham's philosophy that 'the greatest happiness of the greatest number is the foundation of morals and legislation'.[7] But it is an extreme case that implies social welfare increases when total utility is increased, no matter who gets it. The utilitarian social welfare function only considers total utility, and is not concerned with the distribution of utility between generations. It also sets $\varepsilon = 0$, meaning population growth does not increase the discount rate.

At the other extreme, the social welfare function could be min $(u(c_1), u(c_2), ...)$. That is, social well–being is judged by the welfare of the worst off member. This 'perfect egalitarian' case is sometimes known as the Rawlsian social welfare function, after philosopher John Rawls. It considers only the distribution of utility and pays no attention to total utility. If future generations were expected to be better off than the current generation, then the Rawlsian function suggests it should never transfer anything to future generations.

7 Introduction to the Principles of Morals and Legislation (1789).

A more general case would be a social welfare function analogous to power utility:

$W = \sum_{j=1}^{J} \frac{u_j^{1-\varphi}}{1-\varphi}$ where $\varphi \geq 0$ is the constant relative inequality aversion parameter (and setting population growth equal to zero for simplicity). It reflects a value judgement. The higher is φ, the greater the aversion to inequality in utilities. $\varphi = 0$ corresponds to the utilitarian assumption. As $\varphi \to \infty$, we approach the Rawls case. If $\varphi = 1$ then the social welfare function is log linear – or Cobb Douglas. This is the Nash bargaining outcome.

Brennan (2006) argues that the standard social welfare function approach implies when people in a generation are better off, they should receive a lower weight in the social welfare function on equity grounds (extra utility to the better off is not worth as much to society as extra utility to someone worse off). If ongoing economic growth is expected to make future generations better off and if the social welfare function values equality in utility, the social rate of time preference should be positive.

The social rate of time preference is zero with a utilitarian social welfare function, but this an extreme representation of social preferences that is unconcerned with equality between generations.

A positive social rate of time preference does not mean future lives are valued less than current lives. It means that future generations are expected to be better off than the current one, and so extra utility to the future is valued less at the margin than extra utility to the current generation.

But there is another reason for discounting consumption received by future generations. Even if θ, the social rate of time preference is set equal to zero, future consumption is still discounted (the ηg part of the discount rate). The growth in per head consumption means future generations have a lower marginal utility of consumption and so get less from a unit of consumption than earlier generations, so called concavity of the utility function. A dollar to a future generation is less valuable than a present dollar from a social point of view.

Again, this is controversial: for example Kaplow argues that technology may enable future generations to make more effective use of resources to get more out of a unit of consumption.[8]

Although people disagree about the importance of the two reasons for discounting future consumption, there is widespread agreement that future consumption should

8 Kaplow (2008, p. 386).

be discounted.[9] The best way to resolve the disagreements is to express the social welfare function in terms of the consumption of each generation as: $V(c)=\sum_{j=1}^{J}\frac{c_j^{1-\mu}}{(1-\mu)}$ where μ is the relative coefficient of aversion to inequality in consumption.[10] We ignore population issues (again setting population growth $p = 0$).

It is more intuitive to directly think about the effect of the distribution of consumption on social welfare when determining the discount rate for the consumption of future generations. For example, if $\mu = 1$ then we have a log social welfare function, which means 1 per cent of consumption always has the same social value. For example, $200 to someone on $20,000 a year has the same social value as $2,000 to someone on $200,000 a year.

More generally, the contribution to social welfare of a marginal increase in consumption of generation j is $c_j^{-\mu}$. If generation A has k times the consumption of generation B, then the social value of an extra unit of consumption to B is k^{μ} times the value to A. If consumption grows at annual rate g, a generation n years in the future would be $(1+g)^n$ times richer than the current generation and the social value of extra consumption to the current generation is $(1+g)^{n\mu}$ times greater. That is equivalent to discounting future consumption at the rate μg per year (approximately — exact if the variables are expressed in continuously compounded terms).

The coefficient of aversion to inequality in consumption, μ, determines the amount of inefficiency we are willing to bear to pursue redistribution. If generation A has k times the consumption of generation B, then a redistribution from A to B that wasted $1 - k^{-\mu}$ of the transfer would be marginal.[11]

---

9 For example, Arrow et al. (1996), all the contributors in Portney and Weyant (1999a).

10 Note that combining the standard consumer's power utility function and the constant inequality aversion social welfare function does not yield a simple constant elasticity $V(c)$ expression that combines the curvature measures η and φ in the expected way. Although Boadway (2006, p. 3) claims it does, with $\mu=\varphi+\eta+\varphi\eta$, he makes a simple arithmetic error. In fact, substituting the utility function into the social welfare function to express it in terms of consumption means the co-efficient on consumption in $V(c)$ is $(1-\mu) = (1-\varphi)(1-\eta)$, which gives $\mu=\varphi+\eta-\varphi\eta$. But if $\eta > 1$, then $u(c_j) < 0$ creating problems, as in the expression for $W(u)$ it is raised to a real exponent, which is undefined unless the exponent is an integer. Further, a higher φ would reduce rather than increase μ. Kaplow (2003) discusses these problems and the true relationship in more detail.

11 More generally, If $c_i > c_j$ then $SMU_i < SMU_j$ and we transfer from i to j, where SMU is social marginal utility. If making the transfer costs a proportion d of the transfer, we are willing to transfer until $SMU_i = (1 - d)SMU_j$ (ie until the loss to i equals the gain to j). that is, we are willing to put up with waste $d = 1 - SMU_i/SMU_j$. $SMU_i = c_i^{-\mu}$. So are willing to waste $1 - c_i^{-\mu}/c_j^{-\mu} = 1 - (c_j/c_i)^{\mu}$. Note that $(c_j/c_i) < 1$, so the higher μ, the more you are willing to waste, $(c_j/c_i)^{\mu}$ gets smaller.

For example, if μ =1 a redistribution that took $2,000 from someone on $200,000 a year, wasted 90 per cent of it and gave the remaining $200 to someone on $20,000 would keep social welfare constant. If μ =2, a transfer that wasted up to 99 per cent would be worthwhile (for example, a transfer that took $2,000 from A and gave $20 to B would be acceptable).

Harberger (1978) points out this, and other disquieting implications, of the distributional weights approach when considering transfers within a generation and argues that it results in unacceptable outcomes. He concludes that the distributional weights approach does not capture how most people think about distributional issues. It does not represent the value system of most citizens and risks economists' peculiar opinions on distributional issues to swamp all other considerations, something that is beyond the economist's professional role.[12]

In contrast, Stern (2008) and Dietz and Stern (2008) also argue that most people would consider such levels of waste in redistribution undesirable, but conclude that means a μ higher than 2 is implausible, whether arrived at through concavity of the utility function (as in Stern) or through discounting future utility (as in Kaplow and Brennan). They are concerned with action to mitigate global warming. They assume that mitigation would impose a cost on the current generation and benefit much richer future generations. If per capita consumption grows at 1.3 per cent per year (Stern's base case), those living in 100 years time would be 3.6 times richer than people today. Those living in 200 years would be 13.2 times richer.

Their logic is that if it makes no sense to take $2000 off a rich generation to make a poor generation $20 better off, taking $20 off a poor generation to give $2,000 to the rich generation is justified. They assume that the distributional weights approach captures the relevant ethical considerations (although they do agree it is 'a very narrow view of ethics').[13] If μ < 2, then the rate used to discount the future consumption benefits should be no greater than 2g per year (g is the annual growth in consumption) to account for future generations being better off.

A co-efficient of relative inequality in consumption aversion, μ, in the range 1 to 2 would be explained by a coefficient of relative risk aversion η in the range 1 to 2 even with social inequality aversion φ equal to zero. That is, concavity of the utility function in line with standard empirical estimates gives substantial redistribution with all social welfare functions, even a utilitarian one.

Further, although a μ of 2 would be given by $\eta = 2$ or $\varphi = 2$ or some combination, the roles of η and φ in determining μ are not additive. Kaplow (2003) shows that

[12] See Harberger (1978, pp. S118-S119).

[13] Dietz and Stern (2008, p. 104).

the more concave the individual utility functions (the higher is η), the less the effect of the degree of concavity in social welfare (φ) on the co-efficient of relative consumption inequality aversion (μ), especially at high income levels. He concludes 'debates about whether the proper social welfare function is utilitarian or strictly concave (and, if so, how concave) may have diminished practical significance.'[14] The social welfare function may be quite concave in consumption, even if it is not in utility. For example, if η =1, then u_j = $ln(c_j)$ and $W=\sum_{j=1}^{J}\frac{\ln(c_j)^{1-\varphi}}{1-\varphi}=\sum_{j=1}^{J}\ln(c_j)=V(c)=\sum_{j=1}^{J}\frac{c_j^{1-\mu}}{(1-\mu)}$ with $\mu = 1$, independent of φ.

Even with a utilitarian social welfare function, which gives a zero pure rate of time preference ($\theta = 0$), the Ramsey formula is consistent with a wide range of rates to discount consumption (see table 1.3). Suggested consumption growth rates (g) range to 1 to 2 per cent and the co-efficient of relative risk aversion (η) from 1 to 4, which give a discount rate (ηg) of anywhere from 1 to 8 per cent, wide enough to encompass most views on the social discount rate. Not discounting future utility is consistent with substantial discount rates for future consumption.

C.4 Problems with the social welfare function approach

There are a number of problems with this social welfare function approach. First is the lack of agreement on the appropriate function and parameters to choose.[15] How future generations 'should' count is a value judgement and is inherently controversial. There is no consensus about how to value benefits to future generations.

Another problem with the social welfare function approach is that it is totally impractical. To use a social welfare function to do public policy analysis requires measuring utility on a cardinal scale and inter-personal utility comparisons[16] – often assuming that all individuals have the same utility function. If there is to be a trade–off between more happiness for one person, and less for another, we need to be able to measure in a comparable way the changes in happiness accruing to different people. But there is no consensus on whether utility can be measured.

[14] Kaplow (2003, p. 14).

[15] For example, the power utility form is homothetic. Buchanan and Hartley (2000, pp. 135-37) argues that is not appropriate — and the social welfare function should be positively skewed to reflect compassion rather than envy.

[16] Layard and Walters (1978, pp. 45-46).

Inter-generational comparisons are even more difficult. Even if we are happy to assume that the whole current generation can be represented by a single utility function, little is known about the preferences of future generations.

The social welfare function approach is a consequentialist moral theory. It says we should judge policies only in terms of their consequences and the only relevant consequences are individual's gratifications. There is the problem of what weight to attach to preferences that involve envy and maliciousness towards others. It ignores other social goals such as liberty, justice, order, community – goals which transcend individual wants. Further, social choice may be concerned with means.

The social welfare function approach is one particular view of social choice that may not capture how most people think about social welfare or account for equity. For example, the standard form of social welfare function focuses on equality. But most people would be unconcerned about a transfer of income from a very rich person to a rich person, yet a standard social welfare function would say it raises social welfare.

Harberger suggests that, judging by people's charitable giving, redistributions within their family and their gambling behaviour, most seem to care about alleviating poverty rather than equality.[17] He suggests a basic needs approach. Rather than rely on the differential weighting of the welfare of different individuals, this approach imputes external benefits connected with the improvement in the circumstances of others. Most people genuinely believe it is good for the sick to be healed, the homeless sheltered and so on. But it is not the recipient's utility that enters the donor's utility function but the consumption of particular goods and services (food, education, medical care, housing etc) or the attainment of certain states (better nourished, better housed etc) that are closely correlated with the adequate consumption of certain goods and services.

The efficiency approach separates equity and efficiency issues. An advantage is it allows us to explicitly consider the appropriate ethical obligations to future generations. The standard economist's social welfare function approach is only one way to do that. For example, it is not clear that the best way to account for equity effects on future generations is to lower the discount rate used in the social welfare function. Alternative ethical perspectives are possible, which may provide vastly different policy prescriptions.

For example, Nordhaus, suggests:

> Quite another ethical stance would be to hold that each generation should leave at least as much total societal capital (tangible, natural, human, and technological) as it

[17] See Harberger (1984).

inherited. This would admit a wide array of time discount rates. A third alternative would be a Rawlsian perspective that societies should maximize the economic well–being of the poorest generation. The ethical implication of this policy would be that current consumption should *increase* sharply to reflect the projected future improvements in productivity.

Yet another approach would be a precautionary (minimax) principle in which societies maximize the minimum consumption along the riskiest path; this might involve stockpiling vaccines, grain, oil, and water in contemplation of possible plagues and famines. Yet further perspectives would consider ecological values in addition to anthropocentric values.[18]

Some of these decision rules reflect the basic needs approach — people care about ensuring future generations' basic needs are met: not leaving them in dire circumstances and avoiding catastrophes that may threaten society.

C.5 The descriptive approach

The Ramsey formula approach can be interpreted as a positive model of how the economy works. The Ramsey formula then describes the equilibrium market interest rate – the marginal return to capital and the rate at which consumers trade consumption over time. A descriptive approach would then use real world observations to determine the parameters of the model, in contrast to basing them on ethical principles.

One approach is to interpret the utility function of the representative, infinitely–lived household:

$$U = u(c_0) + (1+\pi)^{1-\varepsilon}\gamma u(c_1) + (1+\pi)^{2(1-\varepsilon)}\gamma^2 u(c_2) + ... = \sum_{j=0}^{\infty}(1+\pi)^{j(1-\varepsilon)}\gamma^j u(c_j)$$

as a Barro and Becker dynastic utility function with overlapping generations.[19] In the model parents are altruistic toward their own children. The utility of parents depends on their own consumption, the number of children they have and the utility of each child. Altruistic parents choose family size, consumption and intergenerational transfers by maximising a dynastic utility function.

Altruism justifies the assumption that heads of dynastic families effectively have infinite lives. Because the dynastic head cares about his children's utility, and they care about their children, the head's utility depends on the consumption and number

[18] Nordhaus (2007, pp. 692-93).

[19] See Becker and Barro (1988) and Barro and Becker (1989).

of descendants in all generations. In effect, he acts as if he lives forever and can choose the entire time path of his descendants' consumption. Since the objective function is time consistent, the descendants face a problem of the same form and they have no incentive to deviate from the choices made initially. The dynastic head wants them to maximise their utility, and that is what they do.

The rate of time preference, γ, and the effect of population growth on utility, ε, are now parameters that depend on how the altruism of parents works. That is, the discount rate depends on the preferences of the current generation.

Parents allocate income across bequests, human capital investments in children and their own future consumption so that a marginal dollar gives the same value in all the uses (gives the same marginal contribution to dynastic utility).

Becker and Barro also make fertility and population endogenous, which can dramatically change the model's implications. For example, as the rate used by the current generation to discount the consumption of future generations depends on fertility, it too becomes endogenous.

In the simple infinitely lived consumer models, there are no distortions like public goods, externalities and taxes. The path of consumption determined by utility maximisation and competitive markets is Pareto optimal – one generation cannot be made better off without making some other generation worse off. The social discount rate would be the market rate determined within the model: an efficiency analysis would discount with the market rate of interest. In the presence of distortions, the social discount rate would need to be adjusted, but would still be based on market rates.

If a representative person in the current generation had the same preferences about population and consumption of future generations as a social planner (i.e. if the weights put on future generation's consumption in the social welfare function were the same as in the dynastic utility function), then both would make the same choices for fertility, consumption and investment and both would discount future consumption at the market interest rate.[20]

If the social planner who determines the social welfare function cares more for future generations than does the current generation, the prescriptive approach leads to a lower discount rate than the efficiency based descriptive approach. That is, the social planner places more weight on future numbers of people and consumption than the current generation and would discount the future with a lower rate to maximise social welfare (as judged by the planner's preferences).

20 Barro and Becker (1989).

C.6 Value of life

Some argue that life should not be discounted, that a life saved in the future is no less valuable than a life saved today.[21] Lives cannot be invested and earn interest.

Certainly it is difficult to put a dollar value on a life. Willingness to pay does not work well. For most people, there is no amount that you could pay them to accept immediate, certain death (money is of little use to a corpse). Perhaps some people would accept a large sum knowing it would go to their family — but that puts a lower value on the lives of the most altruistic people. Moreover, the amount most people would pay to avoid immediate certain death would be only limited by their wealth.

Cost benefit analysis does not place a value on human life. Instead, it uses the value of statistical life, which values the reduction in statistical deaths arising from small risks, which is what most government policies affect. It is based on people's observed willingness to pay for small reductions in the risk of death rather than buying out the risk of certain death. For example, suppose a random event, affecting everybody equally, kills one person in a million and that each person is willing to pay $5 to eliminate this risk. A group of one million people is willing to pay $5 million to eliminate the risk of one statistical death to their group. In this example, the value of statistical life is $5 million, which may be greater than each person's wealth. It is $5 divided by the risk reduction of one chance in a million of death.[22] Most government policies are about small reductions in mortality risk and the value of statistical life is the correct way to value the benefits from risk reduction. That is, we are not discounting lives, but the money value of life saving measures. Money can be invested and so the money value of costs and benefits received in the future needs to be discounted.

If regulators valued future lives saved the same as current lives, then it would never be worth spending to save a life today. Money spent today to save lives could instead be invested to produce a larger lifesaving budget in the future, saving more lives. All the more so if technological progress makes the cost of saving lives fall over time. If the value of future lives saved is not discounted, then there is a higher marginal productivity in future spending on lifesaving and all lifesaving resources should be channeled towards the future. But the same argument applies each year. Life-saving expenditures would be delayed indefinitely.[23]

---

[21] For example, Cowen and Parfitt (1992).

[22] Example based on Viscusi (2006, p. 7).

[23] Schelling, (1987).

Choices need to be made between expenditures on reducing the risk to future lives and other goods, not the least of which is saving lives in the present. Putting a money value on the benefits from life saving expenditure just makes explicit what people implicitly do when making choices. People make trade-offs between safety and other uses of resources all the time, for themselves and on behalf of others. People are willing to trade risks to their own lives for quite minor pleasures. Life saving has a financial cost, discounting just allows assessment of the value of expenditures at different periods. 'If willingness to pay to reduce risk is the appropriate metric for allocating regulatory resources, discounting merely adjusts that metric to make expenditures comparable through time'.[24]

[24] Sunstein and Rowell (2007, p. 171).