

23. Addressing Wellbeing in the Long-Term: a Review of Intergenerational Equity and Discount Rates in Climate Change Analysis

Mark Harrison

Discounting and climate change

Most government policies give rise to a stream of costs and benefits over time. To evaluate them requires us to compare costs and benefits received in different time periods. That requires choosing a discount rate, which determines the value of future costs and benefits relative to current ones.

The choice of discount rate can make a significant difference to whether the present value of a project is positive, and to the relative desirability of alternative projects, especially when costs and benefits accrue at different times and over long periods.

A typical project involves upfront costs, with the benefits coming later. If so, the lower the discount rate, the more attractive is the project (the higher its net present value). If the discount rate is set too high, desirable projects may be rejected. If it is set too low, undesirable projects may be approved. The size of the discount rate makes a huge difference to policies where benefits occur in the distant future, such as many environmental policies.

It was a lower discount rate that drove the differences between the policy conclusions of the Stern report and the consensus view of previous cost-benefit analyses of global warming. Stern's cost-benefit analysis of global warming assumed a real discount rate of 1.4 per cent and concluded there was a case for strong action to reduce carbon emissions, recommending an immediate imposition of a high carbon price. Nordhaus assumed a 5.5 per cent discount rate, and favoured a modest carbon price (one-tenth the level of Stern's), increasing over time. The recommended policies differed because of the discount rate assumptions. When Nordhaus ran his computer model using Stern's discount rates, he got similar results to Stern.¹

1 See Nordhaus (2007: 694, 698-700); Baker *et al* (2008: 63-64) and Weisbach and Sunstein (2008: 11).

The difference is not surprising because most of the effects of global warming take place decades in the future. At Stern's discount rate of 1.4 per cent a year, \$1 grows into \$4 in 100 years and \$16 in 200 years. At Nordhaus's 5.5 per cent, it would grow to \$211 in 100 years and \$44,719 in 200 years. Put differently, the present value of \$1 of damages in 100 years with Stern's discount rate is 25 cents, more than 50 times greater than with Nordhaus's. A dollar in 200 years is valued at 6 cents now, almost 2,800 times more than Nordhaus's value.

The Garnaut Report used discount rates of 1.35 per cent and 2.65 per cent.² Yet the average real return on government bonds (both short and long term, nominal and indexed) in Australia was 3-4 per cent over the past 25-40 years. The market return on capital averages 6-10 per cent.³

The mitigation policy to prevent climate change can be viewed as an exhaustible resource problem, which means the discount rate determines the entire path of optimal climate change mitigation. If the objective is to limit atmospheric concentrations of greenhouse gases at some future date, say to 650 CO₂ equivalent ppm, the difference between the limit 650 and current levels (around 390) is an exhaustible resource. The efficient solution is for the marginal costs of emissions reduction and the marginal value of a unit of emissions to rise at the rate of interest used to discount future environmental costs plus the rate at which atmospheric greenhouse gases are re-absorbed. Then the marginal benefit from the emission of a unit to current users is equated to the marginal costs imposed on future generations.

If an ideal carbon price, or cap and trade system, were imposed, the incremental cost of eliminating a unit of emissions and the value of a unit of emissions will be equated to the carbon price, which rises at this rate.⁴

For example, Table 1 shows the price path (ignoring re-absorption) for the carbon price at different discount rates, where the objective is a price of \$300 per tonne CO₂ in a hundred years, say because at \$300, backstop technology that takes greenhouse gases from the atmosphere for recycling or permanent sequestration becomes viable. In practice, the discount rate will affect the time at which we want the price to reach \$300.

2 Garnaut (2008: 19).

3 See Harrison (2010) appendix G.

4 This is derived formally in Becker, Murphy and Topel (2010).

Table 1 Carbon price at different discount rates to reach \$300 per tonne CO₂ in 100 years

		Discount rate			
Year	Years in the future	1.35%	2.65%	4%	6%
2012	0	\$78.48	\$21.94	\$5.94	\$0.88
2032	20	\$102.62	\$37.02	\$13.02	\$2.84
2052	40	\$134.18	\$62.46	\$28.52	\$9.09
2062	50	\$153.44	\$81.13	\$42.21	\$16.29
2072	60	\$175.46	\$105.38	\$62.49	\$29.17
2092	80	\$229.43	\$177.80	\$136.92	\$93.54
2112	100	\$300.00	\$300.00	\$300.00	\$300.00

Source: Author's own work.

If the discount rate is high, it pays to wait and ramp up policy over time. The discount rate makes a huge difference. At the low rates chosen by Garnaut and Stern, the current carbon price is \$78. At Garnaut's higher rate of 2.65 per cent (which Stern now wishes he had chosen) it is less than one-third of that. At four per cent it is less than one-thirteenth. The case for immediate drastic action to mitigate depends crucially on the discount rate chosen. With a six per cent discount rate (a reasonable estimate of the market return) it takes more than 60 years for the carbon price to reach \$30.

The same logic applies in reverse – if the carbon price is set at \$23 and grows at four per cent (as Garnaut recommends), then it would reach \$1,162 in 100 years, or only take 66 years to reach \$300.

Of course this assumes a global carbon price – a partial scheme (such as unilateral action) would reduce the demand for existing fuel sources, bidding their prices down and encouraging greater consumption of existing fuels by those outside the scheme. The net result could be an increase in total emissions if the elasticity of supply of the fuels is low enough (can we prevent existing oil being used?).⁵

Another complication is that for large, non-marginal changes, the interest rate is endogenous. With compounding over long periods, what seem like small amounts can become large. Further, climate change may involve non-marginal effects on productivity and consumption. Policies will involve different paths of interest rates and welfare must be calculated by integrating along each path and comparing the present value of welfare for each policy, using the discount rates consistent with each path. The current rate is the starting point in each case.

The size of the discount rate is a crucial determinant of optimal greenhouse mitigation policy. Those who call for strong immediate action rely on low

⁵ Murphy (2008).

discount rates to justify their policy conclusions – much lower than standard rates used in cost benefit analysis. These low rates rely on strong, controversial assumptions and value judgements, often imposed in a hidden way.

To understand how climate policy modellers derive their low discount rates, we need to consider the fundamentals of economic policy analysis. Often they abandon the traditional cost benefit approach. Instead, climate policy modellers use a social welfare function approach.

The social welfare function approach to policy evaluation

Virtually all public policies involve gainers and losers, affecting lots of people in different ways. Figuring out what changes are desirable when everyone's interests are taken into account is a very complicated problem. Judging whether a change is good or bad requires comparing one person's welfare to another's, a step that requires normative value judgements to be made. Yet any policy decision involves a judgement of its desirability.

One procedure economists use to evaluate outcomes is to use a social welfare function to determine whether social welfare increases. The social welfare function, which Mishan describes as 'that grandiose but ineffectual excrescence on the body of welfare economics'⁶ represents some ethical judgement about the appropriate distribution of welfare across people affected by a policy change. Just as a utility function shows how a person ranks different combinations of consumption goods, the social welfare function represents a value judgement of how society should rank different distributions of utility across people. 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. To bring other factors in is to say social welfare depends on things no one cares about.

As the social welfare function is based on individual utilities, it requires that those functions be measurable and comparable – or cardinal and interpersonally comparable (otherwise the whole notion of going from individual preferences to a ranking of social states flounders on Arrow's impossibility theorem). Many economists would balk at that step – utility cannot be measured, and interpersonal comparisons are difficult to make. Consumer choice theory does not require measurable utility. Utility functions are not unique, any monotonic transformation can be used to describe the same behaviour. Which do we pick? It is somewhat arbitrary.

If there are N individuals in the relevant society, then a simple social welfare function is: $W = w(u(c_1), u(c_2), \dots, u(c_N))$ where c_n is the consumption of person n .

This simple form assumes all individuals have the same utility function (u), which depends only on consumption of a composite commodity (c), representing all goods and services consumed. All are usual, but strong, assumptions that can be generalised (for example, for many environmental problems it would be useful to have utility also depending on an environmental commodity). The utility function is assumed to be concave (declining marginal utility of income). Social welfare is a concave function (w) of individual utilities.

If the social welfare function is, $W = w(u(c_1), u(c_2), \dots, u(c_N))$ the marginal change in social welfare from a change in individual i 's consumption is $\frac{\partial W}{\partial c_i} = \frac{\partial w}{\partial u} \frac{\partial u}{\partial c_i}$ which depends on the w and u functions and the level of c_i . Adding up all the changes in social welfare from the changes in each individual's consumption caused by a policy gives the welfare effects of that policy.

The result is the distributional weights approach. Dollars to low-income people count more highly for social welfare than dollars to high-income people. The change in social welfare for an increase in individual i 's consumption decreases as c_i increases for two reasons: because marginal utility declines as consumption increases (concavity of the utility function decreases $\frac{\partial u}{\partial c_i}$) and because the social weight declines as utility increases.

The key point is how social welfare depends on consumption: the analyst must specify both the w function (how social welfare varies with the utility of each generation) and the u function (how utility of each generation varies with consumption).⁷ If a generation has its consumption increased, the effect on social welfare depends on how that increase affects its utility and then how that affects social welfare. The problem is to maximise w given the individuals are maximising u . The w function is ethical and the u function is empirical. A utility function can be deduced from observing the choices an individual actually makes. But a social welfare function cannot be observed; it must be specified according to a particular ethical view.

As the social welfare function represents some ethical judgement, there is no agreement on the appropriate w function. Economists have no advantage in making ethical judgements, but the social welfare function is useful for making inter-personal utility comparisons explicit and seeing the implications of different value judgements. It allows normative judgements to be introduced in a systematic way.

⁷ This point is strongly made by Kaplow, Moyer and Weisbach (2011).

Comparing the welfare of different generations

Evaluating 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. Evaluating benefits in the distant future requires deriving the appropriate discount rate. The usual approach is to explicitly compare the welfare of different generations with some form of the following social welfare function (based on growth theory models of a representative, infinitely lived household):⁸

$$W = u(c_0) + \frac{(1+p)^{(1-\varepsilon)}}{(1+\theta)} u(c_1) + \frac{(1+p)^{2(1-\varepsilon)}}{(1+\theta)^2} u(c_2) + \dots = \sum_{n=0}^{\infty} \frac{(1+p)^{n(1-\varepsilon)}}{(1+\theta)^n} u(c_n)$$

where each term represents a different generation and c_i is the per person value of lifetime consumption of a typical member of generation i . Note it is assumed each generation can be represented by one utility function and that all generations have the same utility function.

The population grows at the exogenous rate p and θ is the pure social rate of time preference, which is used to discount the utility of future generations (the term ‘pure’ is used for the rate used to discount utility, it is part of the rate used to discount consumption). 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 example, Stern sets it at 0.1 per cent a year (that is an annual probability of one in a thousand).

$(1-\varepsilon)$ is the weighting the social welfare function gives to the number of people in

each generation. For example, if $\varepsilon=1$, then $W = \sum_{n=0}^{\infty} \frac{u(c_n)}{(1+\theta)^n}$ and social welfare depends only on the average consumption of a generation and the number of descendants does not affect the social welfare.⁹ Alternatively, a constant population ($p=0$) would give the same result, and now models each generation as a constant population of identical individuals with the same utility function.

If $\varepsilon=0$, then $W = \sum_{n=0}^{\infty} \frac{(1+p)^n}{(1+\theta)^n} u(c_n)$. This is the ‘Benthamite’ formulation, where the social welfare function maximises the present value of the total utility of all current and future people. Because social welfare depends on both the average utility and number of descendants, when the population is growing,

⁸ See, for example, McCallum (1996).

⁹ For example, Blanchard and Fischer (1989: 38-39) use this case.

the future is more valuable than when only average utility matters (an increase in future consumption per person has a larger effect on social welfare) and social welfare is lower.

Both extremes can give ‘repugnant conclusions’. For example, maximising total utility implies for any population consisting of very good lives there is a better population consisting of a much larger number of people with lives that are barely worth living. Maximising average utility implies for any population consisting of very good lives there is a better population consisting of just one person leading a life at a slightly higher level of wellbeing. More dramatically, the principle also implies that for a population consisting of just one person leading a life at a very negative level of wellbeing, e.g. a life of constant torture, there is another population that is better even though it contains millions of lives at just a slightly less negative level of wellbeing.¹⁰ An intermediate value of ε could avoid these problems.

The standard approach is to assume each generation has the same power utility function $u(c_n) = \frac{c_n^{1-\eta}}{1-\eta}$ with $\eta > 0$, which gives $u(c_j) = \ln(c_j)$ for $\eta = 1$. Power utility has a constant elasticity of marginal utility with respect to consumption, $\eta = -c_j u''(c_j) / u'(c_j)$ (which is also the co-efficient of relative risk aversion).

The production side of the model is assumed to be a standard neo-classical Solow growth model, with exogenous technical progress at rate g . Then, as derived in Appendix 1, if consumption in each generation is chosen to maximise social welfare, in the steady state the equilibrium interest rate is $r = y'(k) - d = (1 + \theta)(1 + p)^\varepsilon (1 + g)^\eta - 1$ where $y'(k) - d$ is the net return on capital and so the rate at which consumption can be transferred from period to period through investing in capital. Consumption per person grows at rate g .

Or approximately (exactly true when expressed as continuously compounded rates, see Appendix 1 for details):

$$r = \theta + \varepsilon p + \eta g$$

This is the Ramsey formula for the social discount rate, accounting for population growth: where θ is the pure social rate of time preference used to discount utility and 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 interest rate r is the appropriate rate to discount consumption. The consumption discount rate (r) is higher than the utility discount rate (θ) because

10 Partridge (2001), Arrhenius, Ryberg, and Tännsjö (2010).

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). Future generations have a lower marginal utility of consumption and get less utility from a unit of consumption than earlier generations, the so-called concavity of the utility function. A dollar to a future generation adds less to wellbeing than a present dollar.

If the marginal rate of return on investment were greater than the Ramsey rate, social welfare would be improved by increasing the capital stock until the equilibrium condition holds.

A positive θ means utility to future generations counts less than utility to the current generation. Some authors argue that, on ethical grounds, θ should be set to zero (or almost zero, with a tiny allowance for the chance of extinction) 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 and the chance of death. Cowen points out 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.¹¹ Further, future generations do not face the risk of death before they are born.

When $\theta = 0$ and $p = 0$ (constant population) the social welfare function has the additive form: $W = u(c_1) + u(c_2) + \dots + u(c_N)$ where social welfare is simply the sum of the individual utilities. This social welfare function is often described as utilitarian, representing Bentham's philosophy that 'the greatest happiness of the greatest number is the foundation of morals and legislation'.¹² Figure 1 illustrates the iso-social welfare curves, analogous to indifference curves, for this social welfare function for a society with two members. An iso-welfare curve shows all the distributions of utility that give the same level of social welfare. The utilitarian social welfare function only considers total utility, and is not concerned with the distribution of utility between people. It requires comparability of changes in utility.

11 Cowen (2001: 5-6).

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

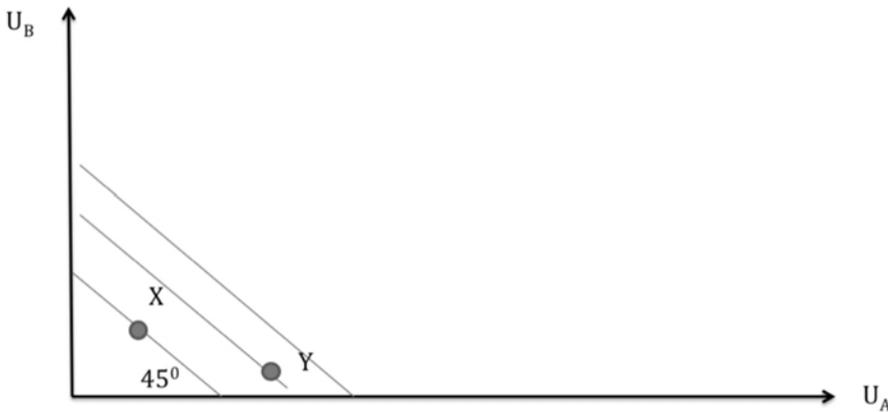


Figure 1 Utilitarian social welfare

Source: Author's own work.

There is fierce debate about whether utilitarianism is an appropriate guide to social welfare. An attractive approach for determining the social welfare function is that individuals should (as an ethical principle) assume they have an equal chance of being each of the people affected by the decision. They will then be in a position to decide objectively whether a policy is in the social interest. For example, in considering what society should look like, consider what society you would like if you were a foetus who did not know your abilities or which part of society you are to be born into (i.e. choosing behind what Rawls, called the 'veil of ignorance' or the original position). Harsanyi (1955) showed that if society's income were fixed and if a person behind the veil of ignorance maximised expected utility, he would want to be born into a society that used a utilitarian social welfare function to determine the distribution of income.

On the other hand, utilitarianism would redistribute income towards 'utility monsters' who derive a great deal of utility and away from people who don't get much enjoyment out of income.

Utilitarianism implies social welfare increases when total utility is increased, no matter who gets it, so the movement from X to Y on Figure 1 increases social welfare, even though it impoverishes person B.

A more general form of the social welfare function, which values both equality and high total social utility, would be: $W = \sum_{n=1}^N \frac{u(c_n)^{1-\phi}}{(1-\phi)}$, where $\phi \geq 0$ is the coefficient of social aversion to relative inequality in utility (a normative

concept), assumed constant in this formulation of the social welfare function (which expresses social welfare as a function of individual utilities in a way analogous to the power utility function). The higher is φ , the greater the aversion to inequality in utilities. $\varphi = 0$ corresponds to the utilitarian assumption and as $\varphi \rightarrow \infty$, we approach the ‘perfect egalitarian’ Rawlsian social welfare function, after philosopher John Rawls where social wellbeing is judged by the welfare of the worst-off member. If $\varphi = 1$ then the social welfare function is log linear – or Cobb Douglas. This is the Nash bargaining outcome.

But even this general form is controversial. Buchanan and Hartley (2000: 135-37) argue that its homotheticity is not appropriate – and the social welfare function should be positively skewed to reflect compassion rather than envy.

How climate change modellers determine the discount rate

The climate change literature uses the reduced form version of the standard social welfare function, expressing social welfare directly in terms of consumption in each generation:

$$W = \sum_{n=0}^{\infty} \frac{c_n^{1-\mu}}{(1+\theta)^n (1-\mu)}$$

where μ is the coefficient of social aversion to relative inequality in consumption, assumed to be constant.¹³ The higher is μ , the greater the aversion to inequality in consumption.

If consumption in each generation is chosen to maximise this social welfare function, in the steady state the equilibrium interest rate is: $r = \theta + \mu g$. The usual approach of climate policy modellers is to determine the discount rate for consumption through specifying these parameters, which are normally expressed in annual terms.

This approach makes strong, often concealed, assumptions.

Most authors ignore the effect of population growth on the discount rate – assuming either constant population ($p=0$) or a Benthamite social welfare function ($\varepsilon=0$). This (usually implicit) assumption gives a lower discount rate (when population growth is positive).

Kaplow, Moyer and Weisbach (2011) point out that the climate modellers approach mixes ethical and empirical parameters and that the determination of the discount rate in climate change models ‘usually proceeds without attention

to the difference between utility functions and the social welfare function and, relatedly, between empirical and ethical discounting'.¹⁴ For example, Stern (2007), Arrow et al (1996), Dasgupta (2008), Nordhaus (2008), Garnaut (2008), Smith (2011) and Quiggin (2006) all use the $r = \theta + \mu g$ formula from the reduced form social welfare function to determine the discount rate for future consumption. But this 'combines empirical facts and ethical views in an opaque and unintuitive manner'.¹⁵

The climate change writers cited above all use the reduced form social welfare function $W = \sum_{n=0}^{\infty} \frac{c_n^{1-\mu}}{(1+\theta)^n (1-\mu)}$ and set $\theta = 0$ on ethical grounds (except Nordhouse, who sets it equal to 1.5 per cent). That makes social welfare the additive form where social welfare is simply the sum of the utility of each generation. That is, they assume the social welfare function is utilitarian. But that is an extreme representation of social preferences that is unconcerned with equality between generations. It assumes the coefficient of social aversion to relative inequality in utility is zero (that is, $\varphi = 0$ in the more general social welfare function form).

At 1.5 per cent annual growth, per capita income in one hundred years will be about 4.5 times the current level. At two per cent the multiple is 7.2. Growth rates in developing countries such as China or India are expected to be much higher.

Brennan (2006) argues that the standard approach implies that when people 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, then $\varphi > 0$ or the social rate of pure time preference, θ , should be positive. This 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.

Further, as Kaplow, Moyer and Weisbach (2011) point out, some modellers suggest that ethical considerations are involved in choosing μ , the coefficient of social aversion to relative inequality in consumption. But this is inconsistent with the modellers' utilitarian assumption that $\theta = 0$, which means social

welfare is just the sum of the generational utility functions, $u(c_n) = \frac{c_n^{1-\eta}}{1-\eta}$. But that means $\mu = \eta$, the elasticity of the marginal utility of consumption, and is entirely an empirical parameter of the utility function.

14 Kaplow, Moyer and Weisbach (2011:9).

15 Kaplow, Moyer and Weisbach (2011:10).

For example, Garnaut supports the ‘utilitarian principle ... that equal weight should be placed on each person’ to reject a positive rate of pure social time preference and then immediately writes:

The second element in the discount rate [i.e. μ] is the marginal elasticity of utility with respect to consumption. This is a measure of society’s concern for equity in income distribution. ... There are compelling theoretical reasons for using an elasticity of 1, ... Dasgupta (2007) argued that an elasticity of 1 implies that ‘distribution of well-being among people doesn’t matter much. ... The Review uses two alternative parameter values for the marginal elasticity of utility, 1 and 2, a range that accommodates strongly diverging views on how much should be spent now to benefit future, presumably richer, generations.¹⁶

But if μ is a measure of society’s concern for equity in income distribution, that contradicts the claim that it is the marginal elasticity of utility with respect to consumption (which is entirely empirical) and that the social welfare function is utilitarian. Once the ethical decision to make the social welfare function utilitarian is made, the issue becomes how consumption affects utility – which is empirical, and nothing to do with further ethical reflection. The utilitarian assumption places no weight on the distribution of wellbeing, it is only concerned about the total. If we reject placing less weight on future generations through a positive θ , why would we want to do it through adjusting μ ? It is not clear what the alternative values of 1 and 2 are supposed to be capturing. One interpretation is that it is changes in predicted behaviour from changes in the utility function, evaluated by a given utilitarian social welfare function. But that is contradicted by the statements that it is about different views on distribution.

The climate policy modellers express the relevant discount rate $r = \theta + \mu g$ in annual terms. As they set $\theta = 0$, the only reason to discount future consumption in their model is because of the declining marginal utility of consumption as generations get richer. A strong implicit assumption has been made, without discussion: the personal pure utility discount rate has been ignored or has been confounded with the social pure utility discount rate, again mixing ethical and empirical parameters.

As the modellers express the discount rate in annual terms, and generations live for more than a year, the model must have overlapping generations. The obvious interpretation is that there is a new generation born each year. Then if $\theta = 0$ and the social welfare function has the additive form: $W = u_1 + u_2 + \dots + u_N$ each term u_i represents the present value, at birth, of generation i ’s lifetime utility.

Suppose generation i is born in period i and lives for T periods and maximises

the power utility function: $u_i = \sum_{t=i}^{T+i} \frac{c_{i,t}^{1-\eta}}{(1+\delta)^{t-i}(1-\eta)}$ where $c_{i,t}$ is generation i 's consumption in period t . Note total consumption in period t is the total of the T generations alive in that period.

$$C_t = \sum_{i=t-T}^t C_{i,t}$$

Then the representative individual would use the discount rate $r = \delta + g\eta$ to discount consumption over a lifetime, where δ is the personal pure rate of time preference, used by an individual to discount his own future utility. It arises from impatience and the chance of death. Crucially, it is empirical, and different from θ , the pure social rate of time preference – which is an ethical concept in a social welfare function used when aggregating different generations' utilities. Empirically, δ is usually estimated at around 1-2 per cent.¹⁷

δ and θ are different concepts and there is no necessary connection between them. It is consistent to believe δ is positive and θ should be set to 0 on the ethical view that all generations should be weighted equally. A person may strongly prefer current to future consumption (high δ) and place a high weight on future generations' consumption (low θ). It would also be consistent to value your own future utility the same as current utility ($\delta = 0$) and not care about affects on future generations after you die (high θ).

The climate policy modellers either ignore the personal discount rate or include it in their θ , used in the reduced form social welfare function and resulting Ramsey formula, and then (without discussion) set it to zero. If so, they mix up the personal and social rate of time preference and their θ 'embodies both empirical and ethical components and cannot be set based solely on either empirical measurements or ethical reflection'.¹⁸

Quiggin (2012) recognises that generations overlap, that θ and δ (using our notation) are distinct parameters and confirms that utilitarianism means $\theta = 0$ – that the utility to all generations counts equally. But he does not consider the appropriate social discount rate and so does not recognise that it may include the personal rate of time preference, and that setting $\theta = 0$ on ethical grounds does not imply that $r = g\eta$.

Quiggin points out that

by the nature of overlapping generations, there is no point at which a coherent distinction between current and future generations can be drawn. In the absence of some general catastrophe, many children alive

17 See Zhuang *et al* (2007: Table 1: 6).

18 Kaplow, Moyer and Weisbach (2011: 10).

today will still be alive in 2100, at which time people already alive will reasonably be able to anticipate the possibility of survival well into the 22nd century.¹⁹

Mishan establishes that if two generations overlap and each receives costs or benefits from a project, then we should discount the costs and benefits with the personal consumption discount rate, $r = \delta + g\eta$.²⁰

For example, assume income is constant over time ($g = 0$) and everyone has a personal pure rate of time preference $\delta = 2$ per cent. Note that the value of an asset with a two per cent return doubles every 35 years, increasing four-fold in 70 years.

The climate policy modeller would not discount future consumption at all, presuming that (with the same utility functions and constant income), \$1 of consumption gives the same utility no matter when it occurs. But if generations receiving the costs and benefits overlap, then a two per cent discount rate should be used. For example, imagine a project in which person A bears a cost of \$100 in 2012. That person dies in 2062. The project gives a benefit of \$500 in 2082 (70 years later) to person B born in 2032. The costs and benefits of the project accrue to generations who overlap (from 2032 to 2062).

Although B has $\delta = 2$ per cent, we cannot say B is indifferent between \$500 in 2082 and \$125 in 2012, this being 20 years prior to his birth. Likewise, we cannot say A is indifferent between these two payments because the \$500 is received 20 years after he dies.

Yet the project should be evaluated with a discount rate of two per cent. For example, evaluate the net value of the project using 2047 (35 years from 2012 and 2082) as our reference year. Person A is indifferent between a payment of \$100 in 2012 and \$200 in 2047. Person B is indifferent between \$500 received in 2082 and \$250 in 2047. Therefore the cost benefit ratio of the project in 2047 is $\$250/\$200 = 1.25 > 1$ and the project is efficient, increases B's utility more than it decreases A's and raises social welfare. The benefits to B exceed the costs to A, whether measured in utility or willingness to pay. Moreover, the cost benefit ratio would be the same in any other reference year common to both lifetimes, both numerator and denominator are changed proportionately by the same discount factor.

Further, the cost benefit ratio is the same as that resulting from the conventional method of calculating net present value in 2012. If we discount back from 2047 to 2012, we reduce both numerator and denominator by a factor of $1.02^{35} = 2$, which gives $125/100 = 1.25$.

¹⁹ Quiggin (2012: 274-275).

²⁰ See Mishan (1981: 196-205).

So if there is a common rate of time preference and an overlap in the lifetimes of people affected by the project, then the discount rate should include the personal pure rate of time preference δ . Whether a project increases or decreases social welfare depends on the changes in utility that different generations experience. If an individual has a positive personal pure time preference rate for consumption, then consumption further in the future adds less to utility than consumption now. A project that pays \$500 in 2082 adds less to B's utility than one that pays \$500 in 2047. A utilitarian social welfare function must accept people's utility functions, including their intertemporal preferences.

Using the personal rate of time preference in the discount rate is not saying that future generations are valued less than current generations – $\theta = 0$ and they count the same. It is merely correctly establishing the effect of each project on each generation's utility, which is an empirical issue, not an ethical one.

If income was growing, then for a utilitarian social welfare function, the discount rate would be $r = \delta + g\eta$, to account for the declining marginal utility of income.

The personal rate of time preference δ is an empirical parameter that is usually estimated as positive (0-3 per cent, usually between 1 and 2),²¹ for good reason (for example, it reflects the possibility of death). Further, the modellers use empirical estimates of η (the elasticity of the marginal utility of consumption), but these are usually jointly estimated with δ . It is not clear why the estimates of η are used, but the same empirical work's estimates of δ are discarded, or if it is legitimate to assume $\delta = 0$ but to use the estimates of η jointly estimated with a positive δ .

Problems arise when generations receiving the project's costs and benefits do not overlap, so that a direct Pareto comparison of their gains and losses is not possible. Their gains and losses can be compounded forward or discounted back while they are alive, but not beyond and so cannot be compared. Now Quiggin may be correct that for most practical purposes, most of those who receive costs and benefits do overlap (but some project beneficiaries may be born after those who pay the costs die).

If there is a time gap between those affected by a project, a potential Pareto improvement is possible if there is some institutional mechanism between generations – such as government transfers, private investment or bequests – that can transform the project's net benefit stream into some new pattern. Then standard discounting (including the personal rate of time preference) shows whether there is a potential Pareto improvement amongst overlapping generations, even if the project costs and benefits are not directly borne by some intermediate generations.

21 See Zhuang *et al* (2007: Table 1: 6).

It is useful to know whether a project covers its opportunity costs. Efficiency is not everything, and a potential Pareto improvement still needs to be judged on whether it improves social welfare – taking account of equity effects. But the efficiency effects of a project are something worth knowing and it is valuable to separate equity and efficiency effects because they are distinct considerations. There is a lack of agreement on how to judge equity effects and we can never expect to achieve a general consensus about the weight that should be attached to the welfare of different groups. Whether the project increases efficiency gives valuable information about the consequences of various options for helping future generations. Efficiency and equity concerns can be separated because we can redistribute across generations not merely through emissions reductions but in many ways, including simply by saving more. Because we can always invest at the market rate rather than in the project being evaluated, any decision to save for the future by investing in a lower-returning project wastes resources. The opportunity cost of one project is other projects foregone.

If anything, efficiency analysis is more useful in an inter-generational context. If the generations are linked through altruism and transfers (including human capital investment parents make in their children and bequests), efficiency gains can make all generations better off. A low return project will have an opportunity cost given by current market interest rates. Examining the efficiency effects of different government projects allows us to maximise the benefit to future generations for a given resource transfer. Even if it is decided to transfer more resources to future generations, the efficiency of particular methods of doing so is still relevant. Also, mixing equity and efficiency risks ignoring the current generation's (possibly offsetting) response to the government project.

Moreover, if the discount rate is to be determined through specifying parameters in the Ramsey equation $r = \theta + \mu g$ on an ethical basis, the wide range in ethical beliefs means it does not resolve disagreement about the appropriate discount rate. Table 2 summarises different parameters that have been suggested. It should be noted that most of the papers referred to fail to clearly distinguish ethical and empirical parameters (such as θ from δ) and set empirical parameters such as η on an ethical basis.

The parameters differ greatly, resulting in recommended discount rates that range from 1.3 to eight per cent. As the differences reflect different judgements about values as well as about the empirical literature, they cannot be resolved objectively. Further, the Ramsey formula gives this wide range even with an assumed zero pure rate of time preference (both θ and δ equal to zero). Suggested consumption growth rates (g) range from one to two per cent and the co-efficient of relative risk aversion (η) from one to four, which give a discount rate (ηg) of anywhere from one to eight per cent, wide enough to encompass most views on the social discount rate.

Table 2 Different discount rates derived from the Ramsey formula

Source	Pure rate of social time preference, θ	Elasticity of marginal utility of consumption, η	Growth rate in consumption, g	Discount rate = $\theta + \eta g$
	per cent		per cent	per cent
Stern (2007)	0.1	1	1.3	1.4
Quiggin (2006)	0	1	1.5	1.5
Cline (1993)	0	1.5	1	1.5
Garnaut (2008)	0	1-2	1.3	1.3-2.6
HM Treasury (2003)	1.5	1	2	3.5
Nordhaus (2007)	1.5	2	2	5.5
Weitzman (2007)	2	2	2	6
Arrow (2007)	0	2-3	If 1-2	2-6
Dasgupta (2006)	0	2-4	If 1-2	2-8
Gollier (2006)	0	2-4	If 1.3	2.6-5.2
Empirical evidence	0-3	0.2-4	1.2-2.1 (for Australia)	0.24-11 (given range)

Source: Author's own work and edited empirical evidence.

Zhuang *et al* (2007) survey empirical estimates of δ (the personal pure rate of time preference, but they do not distinguish it from the social pure rate of time preference) and η (the elasticity of the marginal utility of consumption or the co-efficient of relative risk aversion). They find that δ ranges from 0 to three per cent a year and η from 0.2 to four (with most between one and two).²² Average annual per person consumption growth is usually in the range one to two per cent. For example in recent applications, the Treasury has assumed 1.2 per cent, whereas the Productivity Commission has used 1.75 per cent. Annual GDP growth per capita in Australia averaged 2.1 per cent from 1960 to 2007. Peak to peak labour productivity 1969-70 to 2003-04 averaged 1.71 per cent. Using these parameter ranges in the Ramsey formula gives estimates of a risk-free discount rate ranging from 0.24 to 11 per cent.

Specifying the rate at which society is willing to trade present for future consumption is bound to be controversial. 'Society' is not a decision maker (not even the government controls the whole of society). Inevitably, the analyst imposes a specific discount rate (or the parameters that determine it). Yet economists have no particular expertise about how the future should count.

It is not surprising that there is little agreement about the appropriate social welfare function. What is ethical depends on value judgements and there is no way to reconcile the different value judgements that people may possess.

22 See Zhuang *et al* (2007: 6-7).

Equity issues involve trading off the welfare of one group against another's (for example, the present generation's against future generations') and there is no general consensus about the weight that should be attached to the welfare of different groups. The social welfare function approach makes the discount rate a matter of opinion, and provides no basis for determining which opinion is correct.

Further criticisms of the social welfare function approach

There is not even agreement on whether the social welfare function approach is the appropriate way to judge equity issues. 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 individuals' gratifications. Further, it makes strong assumptions about the form of individual preferences. There is the problem of what weight to attach to preferences that involve envy and malice towards others. It usually ignores other social goals such as liberty, justice, order and community. 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 comfortably rich person, yet a standard social welfare function would say it raises social welfare.

Harberger (1978) points out that the distributional weights approach has a number of disquieting implications. It would only be by accident that the optimal excise tax or subsidy would be zero, as the distributional gains of a change in tax from zero would in general be positive (so long as the gainers and losers had different average incomes). It would generally result in a regressive income tax system. And balancing distributional gains against efficiency losses requires acceptance of large efficiency losses for standard social welfare functions.

For example, in the reduced form social welfare function, assuming $\theta = 0$, the contribution to social welfare of a marginal increase in consumption of generation n is c_n^μ . If $\mu = 1$ then we have a log social welfare function, which means a one per cent increase in 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.

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.

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.²³

For example, if $\mu = 1$ 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 $\mu = 2$, 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) argues this implication of the distributional weights approach results in unacceptable outcomes for transfers within a generation. 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 swamping all other considerations, something that is beyond the economist's professional role.²⁴

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 two is implausible. Action to mitigate global warming would impose a cost on the current generation and benefit much richer future generations. Garnaut and Stern predict much higher income levels in the future – even if the effects of global warming occur. If per capita consumption grows at 1.3 per cent a 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 (and if these levels of growth did not occur, the global warming problem would be very much mitigated).

Their logic is that if it makes no sense to take \$2,000 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').²⁵

23 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 $SMU_i = (1 - d)SMU_j$ (i.e. 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 $1 - c_i^{-\mu}/c_j^\mu = 1 - (c_i/c_j)^\mu$. Note that $(c_i/c_j) < 1$, so the higher μ , the more you are willing to waste, $(c_i/c_j)^\mu$ gets smaller.

24 See Harberger (1978: S118-S119).

25 Dietz and Stern (2008: 104).

As Schelling points out

the 'optimization' approach is based on the principle that if the material benefits we procure for those future high-income people are large enough to offset their reduced marginal utilities, we should procure those future utilities just as if those utilities were our own. Few citizens who understood this principle would ever vote for it. ... we are not used to thinking about making our own sacrifices, or imposing sacrifices on our contemporaries, for the benefit of people who are substantially better off.²⁶

Rather, the willingness to pay current costs to help future generations is associated with potentially catastrophic outcomes of global warming, not from making richer generations even better off.

Harberger suggests that, judging by people's charitable giving and redistributions within their family, most seem to care about alleviating poverty rather than equality. Genuine deprivation – an inability to afford the necessities of life – is what motivates most charity. Further, the altruism we observe is more closely linked to the basic needs of individuals than to their incomes. Most people genuinely believe it is good for the sick to be healed, the homeless sheltered and so on.

Families step in to help fellow members meet basic needs, but seldom redistribute to equalise income (for example, bequests are usually divided equally rather than to offset income differences between children).

Further, the fact that a large proportion of the public participates willingly in lotteries suggests that many people do not value equality. Lotteries increase inequality – many people purchase tickets and make themselves poorer in order to make a few winners rich.

Harberger suggests a basic needs approach.²⁷ 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. 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 externality comes from the extent to which the basic needs of certain segments of society are met. It seems in practice the altruism we observe is more closely linked to the basic needs of individuals rather than to their

26 Schelling (1995: 398).

27 See Harberger (1984).

utility or entire consumption bundle. Citizens want welfare payments to be spent on food and clothing, not beer and drugs. Gifts in kind are preferred to cash transfers.

The basic needs approach accepts these elements of paternalism as a more accurate reflection of our true values, attitudes and beliefs. It is society's views that count, rather than the analyst's. The basic needs approach follows standard cost benefit analysis techniques:

our respective individual gratification (and our underlying willingness to pay) at the fact of a poor child being better nourished is something of a public good. The fact I am gratified in no way precludes you from being gratified also. My willingness to pay to help see that this event comes about can thus appropriately be added to yours, in just the same way that conventional cost benefit analysis adds vertically various citizens' willingness to pay for an environmental improvement that subsequently can be freely enjoyed by all²⁸

Four myths about discounting and climate change

Myth 1: Market discount rates imply we would sacrifice our grandchildren for a few dollars

The only reference to discount rates in Garnaut's 2011 update to the Climate Change Review is on the first page of the introduction where Garnaut recounts how he was explaining to the Multi-Party Climate Change Committee the importance of discounting and

then I said something that brought back the prime minister's attention.

'If we used the share market's discount rate to value the lives of future Australians', I said, 'and if we knew that doing something would give lots of benefits now but would cause the extinction of our species in half a century, the calculations would tell us to do it'²⁹

Economists have calculated the willingness to pay to avoid a catastrophe that wipes out one per cent of the population. Using a six per cent discount rate and assuming income grows at two per cent a year, a permanent one per cent loss in national income in 50 years is worth 25 per cent of income in 2062, which has a present value of 3.52 per cent of current income.

²⁸ Harberger (1984: 112).

²⁹ Garnaut (2011: ix).

Concave $u(c)$ means that reductions in c have rising marginal cost to those who experience them, so that a given reduction in aggregate income is more costly when it is highly concentrated. The costs of a climate-related catastrophe that reduces future national incomes by one per cent by killing off one per cent of the population, while leaving others unharmed, has a much greater cost than a 'marginal' change that reduces everyone's income proportionally. Instead, a portion of the population loses their life – and consumer surplus.

Becker, Murphy and Topel (2010) use the framework provided by the economic literature on the value of a statistical life (VSL), which measures people's willingness to pay for a reduction in the probability of death that would save one 'statistical life'. For example, if in a population of 10,000 individuals each would be willing to pay \$600 a year to reduce the per-annum probability of accidental death by one in 10,000, then $VSL = \$6$ million.³⁰ Drawing on the empirical literature, the value of a life-year is about six times current income. Growth in income over time means future individuals will be willing to pay more to reduce risks to their life compared with poorer current individuals – in 50 years at two per cent annual growth, incomes are 2.7 times greater. The VSL would grow even faster if the value of life were income elastic. In this case, if the discount rate was equal to the rate of income growth, as Stern and Garnaut recommend, the present value of saving future lives would be greater than the value of saving current lives. Murphy and Topel (2006) find the income elasticity of the value of a life-year exceeds 1.33.³¹ On the other hand, a review of sixty studies finds it to be about 0.5 to 0.6.³²

To keep it simple, assume that the value of lives saved from mitigating future catastrophes is proportional to income (an elasticity of one). Then the present value of an event that kills off one per cent of the population in 50 years is at least $6 \times 3.52\% = 21.1$ per cent of current income. Even using market discount rates, future loss of life has a substantial present value.

Willingness to pay does not work well for certain death, as money is no use to a corpse, but clearly willingness to pay to avoid the extinction of all humans in 50 years would be an enormous sum – and so would the amount needed to compensate the current population for that prospect.

30 Becker, Murphy and Topel (2010: 12).

31 Murphy and Topel (2006: 882).

32 See Viscusi *et al* (2005: 723).

Myth 2: Life should not be discounted

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

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. 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.

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.

Cost-benefit analysis converts all costs and benefits to money equivalents based on willingness to pay. The project is exactly equivalent to receiving that flow of dollars over time. The benefit estimates represent the sum of money those who would benefit from the policy are willing to pay to receive the benefits. The cost estimates are the amount of money that needs to be paid to compensate for the costs the project imposes.

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'.³³

Once we express the value in dollar terms, we are saying the benefit is worth that many dollars. Then we must discount, because we can invest dollars to grow over time.

If regulators did not discount the value of future 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 and income growth increases the VSL. 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 channelled towards the future. But the same argument applies each year. Lifesaving expenditures would be delayed indefinitely.

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.

33 Sunstein and Rowell (2007: 171).

Myth 3: High private sector discount rates are specific to the share market

Garnaut also informed the committee that

So we had to choose the right discount rate', I said. 'We can't use the discount rates that determine values in the share market, because they take into account risks of a kind that are not relevant here.'³⁴

The opportunity cost relevant for climate change mitigation is the foregone return on equivalent private sector investment. Direct measurement of these returns shows them to be high. For example, using national accounts data to divide the total income from capital generated in the private sector by an estimate of the private sector capital stock shows real returns to capital to be consistently high, a number of studies showing they average over eight per cent real for decades, and are much more stable than share market returns (with a 95 per cent confidence interval of plus or minus one percentage point, whereas it is plus or minus four percentage points for share market returns). These estimates are consistent with share market returns, which are after company and property taxes. Estimates of the cost of capital in Australia and of the return to human capital investment are also high.³⁵ Foregone private investment has a high opportunity cost which is not limited to the share market. In fact, capital market imperfections, such as credit rationing, high borrowing rates and imperfect annuity markets, are likely to make the returns to non-listed companies and family investment in human capital higher than for listed companies.

The high private sector returns relative to the risk-free return (usually measured by the return on government bonds) indicate high-risk premiums are not limited to the share market, but seem to reflect broader attitudes to the cost of risk.

The government should only invest in a project if it can reap a better return from it than from investing in a private sector project of equivalent risk. That requires using the private rate of return to capital as the discount rate.

Myth 4: Governments should discount with the risk-free rate of return because of the Arrow and Lind theorem

The climate policy modellers invariably discount with a risk-free rate. For example, they use the Ramsey equation for the equilibrium risk-free return to derive the social discount rate.

³⁴ Garnaut (2011: ix).

³⁵ See Harrison (2010: 37-39, 128-34).

In a widely cited, classic contribution to public sector discounting theory, Arrow and Lind (1970) showed that if a government project was ‘small’ (in relation to the total wealth of taxpayers) and ‘the returns from a given public investment are independent of other components of national income’, then the social cost of the risk for project flows that accrue to taxpayers tends to zero as the number of taxpayers tends to infinity.³⁶ That is, government investments with diversifiable risks spread over many households should be evaluated using the riskless rate to discount expected benefits (that is, with no adjustment for risk).

This result, known as the Arrow-Lind theorem, is consistent with the CAPM approach. If a project contains only diversifiable risk and no aggregate risk, an efficient private sector would spread the risk and would also use the riskless rate to discount expected project returns.

If a project contains aggregate risk, then a risk premium should be used. Aggregate risk is an irreducible social risk that cannot be diversified, even by government. It is caused by shocks such as recessions and variations in the market return. Most government projects involve aggregate risk. As Bailey and Jensen point out:

the ‘private’ (and ‘social’) risk of even a small project which is perfectly correlated with the average returns on all other assets cannot be reduced one iota by transferring it from the private to the public sector.

The question regarding the size and sign of the covariances of returns on prospective projects is an empirical issue. However, some brief consideration of the problem seems to indicate (contrary to Samuelson et al) that the vast majority of government projects will have outcomes correlated with national income. For instance, any government investment that facilitates ordinary commerce will produce more benefits when national income is high than when it is low. Electric power, highways, waterways, airports, and postal service, for example, all have this character.³⁷

The government should only price risk differently if it has some advantage that allows it to improve on an imperfect market. The advantage should be specified, because the government cannot correct some market imperfections. Any comparison of how the government and private sector manage risk should take account of how the government in fact operates under the incentives of the political process.

If governments do not spread the risk, then it is borne by individuals and discounting should be done at a rate that reflects the risk premiums they demand for bearing risk. Arrow and Lind conclude

³⁶ Arrow and Lind (1970: 171).

³⁷ Jensen and Bailey (1972: 7).

some benefits and costs of sizeable magnitudes accrued directly to individuals so that these individuals incurred the attendant costs of risk bearing ... it is appropriate to discount for risk ... As a practical matter, Hirshliefer's suggestion of finding the marginal rate of return on assets with similar payoffs in the private sector, and using this as the rate of discount, appears reasonable for discounting those costs and benefits which accrue privately.³⁸

The market price of risk shows the cost of extra non-diversifiable risk to the private sector, whether the market is efficient in handling risk or not.

The evaluation of the benefits from climate change mitigation should take account of the great uncertainty about the future costs and benefits of such policies, such as uncertainties about the distributions and size of climate damage and the payoffs from mitigation. Becker, Murphy and Topel (2010) set out the appropriate framework to do so. For example, the present value of mitigation increases greatly when climate-related damages are unequally distributed, when future lives are at risk, and when we add uncertainty as to when the damaging events might occur (holding constant the expected time of occurrence).

Risk should be taken into account in the standard way – by adjusting the market interest rate used to discount future flows – not by asserting that a risk-free rate should be used. This allows us to compare how policies deal with the risks; some policies become more attractive precisely because they pay off in bad times. By explicitly accounting for risk, we can distinguish between policies that reduce risk and those that don't.

The appropriate adjustment to the discount rate depends on the covariance of the returns to mitigation with m , the marginal rate of substitution between present and future consumption (which depends on both income growth and the state of the environment, falling if income is growing and rising if the environment becomes degraded). If we divide by the variance in m , we get the environmental project's 'beta'.

If the environmental asset offers greater payoff than the market when m is high (say because the environment is degraded), then it reduces risk and the discount rate would be below the market return (beta is less than one).

If the returns were positively correlated with m , then the expected return would be below the risk-free return (beta is negative). If the returns on mitigation investments avert disasters and avoid large reductions in productivity and living standards or save lives, they pay off exactly when other assets do not and when willingness to pay is great. The appropriate discount rate is below

38 Arrow and Lind (1970: 175-76).

the risk-free rate. That is, when climate policies effectively insure against large downside risks, they have low expected return but high market value because they pay off when the mitigation of damage is most valuable.

That is not true for modest climate change – not costly enough to reduce consumption (both Stern and Garnaut predict growing consumption even with climate change). The gains from mitigation against modest climate change are greatest in the highest GDP states and so have a positive beta.³⁹ For example, where losses are proportional to GDP, the appropriate beta is one and the discount rate should be the market return. Policies that raise future GDP, already 5-8 times higher than current level, by a few per cent are not that valuable.

Adjusting appropriately for risk, rather than pretending it is irrelevant, allows us to judge the best policies – those with the greatest net present value, such as policies that focus on avoiding extreme outcomes that could happen quickly. For example, Becker, Murphy and Topel (2010) emphasise that research and development investments in mitigation technologies that can be scaled up in the event that damages are large, can offer important insurance against looming catastrophe and should not be heavily discounted. Economically useful policies would encourage research and development into finding such scalable technologies. Further, investments in technology that provide future generations with the means to cope with global warming are less reversible and avoid the problem of ensuring future intervening generations commit to the plan.

Conclusions

A number of assumptions and value judgements go into the claim that discount rates of one to two per cent should be used to evaluate mitigation policies. The case for immediate mitigation to prevent climate change depends crucially on the use of low discount rates, yet the climate policy modellers' standard approach jumbles ethical judgements and empirical facts and makes concealed presumptions that many people would disagree with.

Good cost benefit analysis makes clear and transparent the assumptions and judgements made, so that they can be scrutinised and debated. By contrast, the Stern Review did not even reveal its discount rate in the initial report and as Cowen points out 'the relevant caveats don't seem to find their way into his final presentation of the estimates'.⁴⁰

It is important to separate positive analysis from normative beliefs, because there is less agreement about ethical judgements. At the least, it helps clarify the effects of policies to clearly distinguish empirical parameters that determine

³⁹ See Murphy (2008).

⁴⁰ http://marginalrevolution.com/marginalrevolution/2006/11/the_stern_repor.html.

behaviour from ethical judgements about how to evaluate results. Further, the wide range of ethical beliefs means the social welfare function approach does not even resolve the issue of the appropriate discount rate – but can support a wide range of risk-free social discount rates from one to 11 per cent.

The ten commandments followed when climate policy modellers choose discount rates are, that:

1. The social welfare function approach is the only relevant way to think about the outcomes of climate change policy.
2. The appropriate social welfare function is utilitarian ($\varphi = 0$) and future utility should not be discounted ($\theta = 0$), which presumes equity between generations is irrelevant.
3. The fact that individuals discount their own future utility is irrelevant ($\delta = 0$).
4. We should maximise total utility ($\varepsilon = 0$) so that population growth is irrelevant for the discount rate.
5. The ethical discount rate is relevant and the current population and future generations will support efforts to direct more to generations still further in the future.
6. Efforts to help future generations won't be offset by the current generation's behaviour.
7. The Ramsey rule is the correct positive model of interest rate determination, with no regard to whether the evidence supports it or alternative models.
8. Risk is irrelevant; the risk-free rate should be used.
9. Despite all these assumptions, there is no need to conduct sensitivity analysis on the discount rate.
10. Whether the transfer is efficient is irrelevant.

The most significant presumptions are the departure from well-established and justified cost benefit analysis in the final two commandments, especially the abandonment of standard efficiency based analysis. Instead, climate policy modellers adopt what Nordhaus calls

the lofty vantage point of the world social planner, perhaps stoking the dying embers of the British Empire, in determining the way the world should combat the dangers of global warming. The world, according to Government House utilitarianism, should use the combination of time discounting and consumption elasticity that the *Review's* authors find persuasive from their ethical vantage point.⁴¹

Appendix 1: Deriving the Ramsey rule

The social welfare function is

$$W = u(c_0) + \frac{(1+p)^{1-\varepsilon}}{(1+\theta)} u(c_1) + \frac{(1+p)^{2(1-\varepsilon)}}{(1+\theta)^2} u(c_2) + \dots = \sum_{n=0}^{\infty} \frac{(1+p)^{n(1-\varepsilon)}}{(1+\theta)^n} u(c_n)$$

where each period represents a different generation and c_n is the per person consumption of a typical member of generation n . The population grows at the exogenous rate p and $\theta > 0$ is the pure social rate of time preference, which is used to discount the utility of future generations. $(1-\varepsilon)$ is the weighting the social welfare function gives to the number of people in each generation.

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 generation is a constant returns function of per household capital K and efficiency units of labour, L : $Y_j = Y(K_n, L_n)$ where $L_n = (1+g)^n (1+p)^n$. 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 $y_n = Y_n/L_n$ and $k_n = K_n/L_n$ and

$$y_n = (1/L)Y(K/L, 1) = y(k).$$

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, $Y_K = \partial F/\partial K = \partial f/\partial k$.⁴² 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.⁴³

In the steady state, output, consumption, the wage and capital per efficiency unit are constant. But it is people, not efficiency units, who 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 possible only if u' has a constant elasticity with respect to c_n . Assume, therefore, that each representative

household has a power utility function: $u(c_n) = \frac{c_n^{1-\eta}}{1-\eta}$ with $\eta > 0$, which gives $u(c_n) = \ln(c_n)$ 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.

⁴² $\partial F/\partial K = \partial(Nf)/\partial K = N(\partial k/\partial K)(\partial f/\partial k) = \partial f/\partial k$.

⁴³ In the case of a Cobb-Douglas production function, Harrod neutral technical progress is also Hicks neutral. $Y = K^\alpha [a(t)L]^{1-\alpha} = [a(t)]^{1-\alpha} K^\alpha L^{1-\alpha}$

The household's budget constraint is:

$$y(k_n) = c_n + (1 + p)k_{n+1} - (1 - d)k_n$$

where capital depreciates at rate d .

Choose values of c_n and K_1, K_2, \dots to maximise social welfare:

$$W = u(c_0) + \frac{(1+p)^{(1-\varepsilon)}}{(1+\theta)} u(c_1) + \frac{(1+p)^{2(1-\varepsilon)}}{(1+\theta)^2} u(c_2) + \dots = \sum_{n=0}^{\infty} \frac{(1+p)^{n(1-\varepsilon)}}{(1+\theta)^n} u(c_n)$$

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):

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

If a variable has a growth rate of x per period, then it is $1+x$ after one period. If it is continuously compounded at rate x^* , then it grows to e^{x^*} after one period. So a growth rate of x is equivalent to a continuously compounded rate of x^* when $e^{x^*} = 1 + x$. Taking logs gives $x^* = \ln(1+x)$.

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.

References

- Arrhenius, Gustaf, Ryberg, Jesper and Tännsjö, Torbjörn, 'The Repugnant Conclusion', *The Stanford Encyclopedia of Philosophy (Fall 2010 Edition)*, Edward N Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2010/entries/repugnant-conclusion/>>.
- Arrow, K, M Cropper, G Eads, R Hahn, L Lave, R Noll, P Portney, M Russell, R Schmalensee, V Smith, and R Stavins (1996). *Benefit-Cost Analysis in Environmental, Health, and Safety Regulation* American Enterprise Institute, The Annapolis Center, and Resources for the Future.
- Arrow, K and R Lind (1970). 'Uncertainty and the Evaluation of Public Investment Decisions' *American Economic Review*, v.60, no.3: 364–78. Reprinted in Layard and Glaister (eds) (1994): 160–178.
- Becker, G, K Murphy and R Topel (2010). 'On the Economics of Climate Policy' *George J Stigler Center for the Study of the Economy and the State Working Paper no.234*, June.
- Blanchard, O, and S Fischer (1989). *Lectures on Macroeconomics*. Cambridge, Mass.: MIT Press.
- Brennan, G (2006). 'Discounting the future, yet again' seminar paper for the Research Triangle Ethics Circle. <http://www.unc.edu/~gsmunc/EthicsGroup/Brennan%20--%20Discounting%20the%20Future.pdf>
- Cowen, T (2001). 'What is the Correct Intergenerational Discount Rate?' Mimeo. Department of Economics, George Mason University.
- Dasgupta, P (2006). 'Comments on the Stern Review's Economics of Climate Change' December 12. Seminar on the Stern Review's Economics of Climate Change, organised by the Foundation for Science and Technology at the Royal Society, London,
- Dietz, S and N Stern (2008). 'Why Economic Analysis Supports Strong Action on Climate Change: A Response to the Stern Review's Critics' *Review of Environmental Economics and Policy*, vol.2, no.1, Winter: 94-113.
- Garnaut, R (2011). *The Garnaut Review 2011, Australia in the Global Response to Climate Change* Cambridge University Press.
- _____ (2008). *The Garnaut Climate Change Review*. Final report. Cambridge University Press.
- Harberger, A (1984). 'Basic needs versus distributional weights in social cost-benefit analysis', *Economic Development and Cultural Change*, vol 32, no 3 (April): 455–474.

- ____ (1978). 'On the Use of Distributional Weights in Social Cost-Benefit Analysis' *Journal of Political Economy*, vol.86, no.2, part 2, April: S87-S120
- Harrison, M (2010). 'Valuing the Future: the social discount rate in cost-benefit analysis' *Visiting Researcher Paper, Productivity Commission*, Canberra. <http://www.pc.gov.au/research/visiting-researcher/cost-benefit-discount>
- Harsanyi, J (1955). 'Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility' *Journal of Political Economy*, v.63, no. 4, August: 309-321.
- Jensen, M and M Bailey (1972). 'Risk and the Discount Rate for Public Investment' in *Studies in the Theory of Capital Markets*, M. Jensen, (Ed), Praeger Publishers. <http://papers.ssrn.com/abstract=390110>.
- Kaplow, L (2006). 'Discounting Dollars, Discounting Lives: Intergenerational Distributive Justice and Efficiency' *Prepared for the University of Chicago Conference on Intergenerational Equity and Discounting*, April 10.
- Kaplow, L, E Moyer and D Weisbach (2011). 'The Social Evaluation on Intergenerational Policies and Its Application to Integrated Assessment Models of Climate Change' *John M Olin Law & Economics Working Paper No. 565, The Law School, The University Of Chicago*, July
- ____ (2003). 'Concavity of Utility, Concavity of Welfare, and Redistribution of Income' *Harvard Law School John M Olin Center for Law, Economics and Business Discussion Paper Series*, Paper 437.
- McCallum, B (1996). 'Neoclassical vs. Endogenous Growth Analysis: An Overview' *Federal Reserve Bank of Richmond Economic Quarterly*, Vol. 82, no.4 Fall: 41-71.
- Mishan, E (1981). *Economic Efficiency and Social Welfare*, George Allen and Unwin, London.
- Murphy, K 'Some simple economics of climate change' Presentation to Mont Pelerin Society 60 the Anniversary Conference, Tokyo, September 7.
- Murphy, K and R Topel (2006). 'The Value of Health and Longevity' *Journal of Political Economy*, 2006, vol. 114, no. 5: 871-904.
- Nordhaus, W (2007). 'A Review of the Stern Review on the Economics of Climate Change' *Journal of Economic Literature*, Vol. XLV (September 2007): 686-702.
- ____ (2007b). *The Challenge of Global Warming: Economic Models and Environmental Policy*.

- Partridge, E (2001). 'Future Generations' from *A Companion to Environmental Ethics*,
 Edited by Dale Jamieson, Blackwell. <http://gadfly.igc.org/papers/futgens.htm>
- Portney, P and J Weyant (1999). 'Introduction': 1-11 in Portney and Weyant (1999a).
- Quiggin, J (2006). 'Stern and the critics on discounting'.
- ____ (2012). 'Equity between overlapping generations' *Journal of Public Economic Theory*, 14 (2): 273–283.
- Schelling, T (1995). 'Intergenerational discounting' *Energy Policy*, v.23, no. 4/5: 395-401.
- Smith, K (2011). 'Discounting, risk and uncertainty in economic appraisals of climate change policy: comparing Nordhaus, Garnaut and Stern' Department of Climate Change and Energy Efficiency, 3 February.
- Stern, N (2008). 'The Economics of Climate Change' *American Economic Review: Papers & Proceedings*, vol.98, no.2, May: 1–37.
- ____ (2007). *The Economics of Climate Change: The Stern Review*, Cabinet Office – HM Treasury, Cambridge University Press, UK.
- Viscusi, W, J Harrington and J Vernon (2005). *Economics of regulation and anti-trust*. MIT Press.
- Weisbach, D and C Sunstein (2008). 'Climate Change and Discounting the Future: A guide for the Perplexed' *Reg-Markets Center Working Paper no.08-19*, *University of Harvard Public Law Working Paper no.08-20*.
- Zhuang, J, Z Liang, T Lin, and F De Guzman (2007). 'Theory and Practice in the Choice of Social Discount Rate for Cost–benefit Analysis: A Survey' *ERD Working Paper No. 94*, Asia Development Bank, May.