This version: 20 Dec 2006

Stern and the critics on discounting

John Quiggin

Australian Research Council Federation Fellow

School of Economics and School of Political Science and International Studies

University of Queensland

EMAIL j.quiggin@uq.edu.au

PHONE + 61 7 3346 9646

FAX +61 7 3365 7299

http://www.uq.edu.au/economics/johnquiggin

I thank Nancy Wallace for helpful comments and criticism.

This research was supported by an Australian Research Council Federation Fellowship.

Stern and the critics on discounting

The Stern Review of the economics of climate change has had a substantial impact on public debate over policy responses to climate change. This is in part a matter of timing. Although economists have debated the relevant issues for years, this debate has been overshadowed by debate over scientific issues, initially among climate scientists and then, for much of the last decade, between scientists and political critics of climate science, primarily associated with US-based think tanks . The operations of these groups in relation to climate science and other environmental issues, evolution and 'intelligent design' theory, stem cell research and other issues is discussed in detail in Mooney (2006)

By 2006, the credibility of the anti-science groups was on the verge of collapse, at least as regards climate change. Factors contributing to this outcome included the resolution of remaining scientific controversies, particularly relating to satellite measurements of tropospheric temperatures, the success of the documentary *An Inconvenient Truth* and increased publicity regarding the extent to which prominent skeptics were beneficiaries of funding from fossil fuel industries.[1]

The Stern Review radically changed the terms of the debate by presenting the issues in economic rather than scientific terms. The effects of global warming, previously discussed in qualitative terms were shown to correspond to large losses in economic welfare. The result was to outflank the remaining skeptics. They could either continue denying the results of scientific analysis, or try to salvage the fallback position, undermined by the Stern Review, that although global warming is real, the costs of doing anything significant about it exceed the benefits, at least in the short term. It is this latter position which will be critically examined in the present paper.

[1] Even more damaging was the disclosure that several prominent skeptics had earlier worked for the tobacco industry, challenging scientific analysis of the dangers of passive smoking.

Some aspects of the Stern Review have been accepted with little controversy. Most significantly, the estimate that stabilising global CO2 levels at 500 ppm woud reduce world income by around 1 per cent has been generally accepted as reasonable, if towards the optimistic end of the range of plausible estimates. Although most serious economic analysis has produced cost estimates for climate stabilisation of 1 to 2 per cent of world income, much informal discussion has presupposed far larger estimates.

There has also been little discussion of the Stern Review's projections for climate change, which generally followed those of the IPCC or of estimates of the economic value of damage to natural ecosystems which broadly followed those of Nordhaus and Boyer. Nordhaus and Boyer have been criticised for undervaluing damage to natural environments (Quiggin 2006) and it seems likely that a similar undervaluation will be evident in Stern.

The main focus of discussion of the Stern review has been the way in which future costs and benefits. This review will focus on the same question. It begins with an outline of the expected utility model used in the Stern review, then examines 'pure' time discounting before considering the Stern review and its critics.

Expected utility

Stern, and nearly everyone else in the debate so far, uses a model based on expected utility theory. There are very strong reasons to go this way. First, expected utility has the property of dynamic consistency, which means that, if you make a plan, anticipating all possible contingencies, you will want to continue with that plan over time, whichever contingency arises. No other choice model has this property except under special conditions.

Second, expected utility theory allows a single utility function that simultaneously determines attitudes to intertemporal wealth transfers, interpersonal redistribution and risk reduction (transfers of income between states of nature). With the plausible technical assumption of constant relative risk aversion, (almost) everything is determined by a single parameter (called eta in the Stern report), which measures the elasticity of substitution for income.

The central idea of expected utility theory is that people are not concerned ultimately with money income but with the utility derived from consumption. The assumption of diminishing marginal utility means that a dollar of extra income is worth less if you're rich than if you're poor. So we derive the conclusions

(i) assuming rising incomes, a dollar of extra income is worth less in the future than it is today

(ii) under uncertainty, a dollar of extra income in a bad state of nature is worth more than a dollar in a good state of nature

(iii) transferring income from rich people to poor people improves aggregate welfare

The first of these assumptions mean that in the presence of technological progress, allowing rising incomes, we would expect a positive discount rate. The second means that a risky asset (more precisely, a risky asset with returns that are correlated with aggregate consumption) should be worth less than a riskless asset. The third means that even when redistributive taxes and international aid are costly (for practical purposes, nearly always) they can improve welfare.

Assuming that the combination of the expected utility model and inherent discounting (discussed below) captures all the issues under consideration, the riskless discount rate is determined by a simple formula

$$r = delta + eta*g$$

where eta is the elasticity of substitution for consumption, g is the rate of growth of consumption per person, and delta is the inherent discount rate. A similar, slightly more complex formula can be used to derive the rate of return for a risky asset, based on its correlation with aggregate consumption.

The parameter eta represents the extent to which the marginal utility of income is reduced as income increases, and the choice of eta is central to the debate over discounting. The larger is eta, the stronger all of effects (i)-(iii) become. So high eta means a high preference for current income, high aversion to risk and large benefits from redistribution.

Even economists familiar with the mathematical derivation of eta often have problems understanding the implications of different choices of eta, particularly when time, uncertainty and interpersonal redistribution interact. So it may be useful to consider a particular example, which happens to be that used in the Stern Review.

The special case eta = 1

Expected utility theory was first developed to analyze choice under uncertainty. In discussions focused on risk, the most common single choice for eta is eta=1, corresponding to a logarithmic utility function. This is a particularly tractable function, and seems to fit the data reasonably well. It also gains popularity from tradition having been proposed as a utility function for money by Daniel Bernoulli back in 1738.

There is a natural way of interpreting eta=1, that is, logarithmic utility, in the intertemporal context. With this specification (and ignoring inherent discounting as discussed bleow) one per cent of income now has the same value as one per cent of income at any time in the future. So, for example, a policy that made income (not the growth rate of income!) by one percentage point from 2000 to 2050, relative to some baseline, then increased income by one per cent relative to the same baseline, until 2100 would come out exactly neutral. Logarithmic utility is implicit in much discussion of intergenerational equity, notably the intergenerational accounting analysis popularized by Kotlikoff and Summers, which focuses on the proportion of income paid in tax by each generation.

At this point a numerical example might be useful. The world's mean income per person is currently around $US7000, though the majority of people get much less and the billion or so in rich countries much more. Under the projections used in the Stern Review, average world income in 2100 is estimated at about $US 100 000. Using eta =1, a sacrifice of $70 per person (1 per cent of income) today would be justified if (and only if) it increased the income of our great-grandchildren in 2100 by at least $1 000. If this trade-off appears reasonable, then a value of eta =1 is appropriate. If the future payoff required is higher (or lower) then so is the preferred value of eta.

Intuition about the far future tends to be cloudy, so it is worth observing that, under expected utility theory, exactly the same arguments apply to redistribution within the current generation. To illustrate, it is useful to turn around the direction of redistribution. Consider a redistributive program that takes $1000 from the well-off (in this example households with income of $100 000 per person) and uses the proceeds to benefit the poor (those with $7000 per person). (Alternatively to keep the focus on redistribution from the poor to the rich, suppose that such a program already exists, and consider scrapping it).

Such redistribution always involves a range of costs including administration, compliance, efforts at avoidance and evasion and incentive costs. Suppose that, in a particular case, these costs amount to 93 cents per dollar initially taxed, so that for a net loss of $1000 to the well-off, the net benefit to the poor is $70. If such a program is exactly marginal, so that any program with a larger net benefit is acceptable, and any program with a smaller net benefit is unacceptable, then the implied social preferences have eta = 1. If the minimum acceptable net benefit is larger (smaller) then we require eta to be less than (greater than one).

The choice of eta =1 also seems to give a good match for calibrated macroeconomic models designed to match growth and business cycle facts. On the other hand, studies of aggregate consumption seem to imply values of eta close to zero, while studies of asset prices for equity seem to imply values well above one (as discussed below in relation to the equity premium puzzle)

Problems with expected utility

The expected utility model is neat, logically compelling and tractable, but it suffers from two big problems.

First, at least some of the time, most people don't behave in a way that is consistent with the expected utility model. For example, people simultaneously gamble at unfavorable odds and take out costly insurance, which violates the predictions of EU with respect to uncertainty. Similarly people often apply a high discount rate to trades between the present and the near future, but a low discount rate for trades between the near and far future. This is called hyperbolic discounting. Large branches of modern economic theory, commonly referred to as generalized expected utility theory and behavioral economics, attempt to address this problem as discussed below.

The second problem is that observed market outcomes are not consistent with EU theory as it is commonly applied. This problem is partly because people don't act in accordance with EU and partly because markets don't work in the smooth and frictionless way assumed in standard finance-theory models.

The most important problem in this respect is the 'equity premium puzzle', and the closely-related 'risk-free rate puzzle'. The equity premium puzzle is that for plausible choices of eta, the real bond rate should be somewhat higher than it is, and the rate of return to equity much lower.

Historically, real returns to investors from the purchases of U.S. government bonds have been estimated at one percent per year, while real returns from stock ("equity") in U.S. companies have been estimated at seven percent per year, a difference of six percentage points. By contrast, for reasonable choices of eta, the difference should be no more than half a percentage point. The equity premium puzzle can be resolved by assuming very high values of eta since risk aversion increases the premium. But high values of eta imply a high discount rate, so the risk-free rate puzzle is made worse.

The inconsistencies between the EU model and observed choices and market outcomes mean that for any possible choice of parameters, it's possible to present hypothetical choices for which most people will reject the implications of the model, or to point to market outcomes inconsistent with the proposed parameters.

Inherent discounting and weighting

Although expected utility provides a complete theory of income allocation across individuals, time periods and states of nature, it is often supplemented by some sort of weighting scheme. This is true whether expected utility is used positively, to model actual behavior, or normatively as a guide to rational individual decision-making and ethical social-decision-making.

In the case of allocating consumption over time we need to consider whether we should discount future income simply because it is in the future, even with the same marginal utility (call this inherent discounting). Similar issues arise in relation to the weighting of individuals in income redistribution and states of nature in choice under uncertainty, and these will be considered first.

Probability weighting and uncertainty

In decisions under uncertainty, individuals often seem to put more weight on low-probability extreme outcomes than would be implied by expected utility (Kahneman and Tversky 1979). Although various responses to this observation have been proposed, the most popular has been to use a rank-dependent weighting scheme, as proposed by Quiggin (1981, 1982, 1993) and incorporated in the cumulative prospect theory model of Kahneman and Tversky (1992).

Application of rank-dependent weighting schemes is controversial in a normative sense, since it leads to violations of dynamic consistency, except where decision-making starts from a particular point and holds relative weights fixed thereafter. On the other hand, if individuals make decisions using probability weights, it is necessary to incorporate this fact in public policy. For example, if individuals overweight the possibility of a nuclear power-plant accident, and take expensive precautions against it, the construction of the plant may reduce welfare, even if an expected-utility calculation, assuming that everyone followed expected-utility, would yield a net benefit.

Interpersonal weighting

The idea that all individuals' utility should have equal weight is clearly not valid as part of a descriptive model if individual behaviour and has been controversial as a normative basis for social welfare. However, the deviations from equal weighting tend to go in opposite directions.

As individuals, we almost invariably place more weight on our own welfare and that of our immediate family than on the welfare of friends, neighbours and more distant relations. In turn, we value the welfare of fellow-citizens in general less than that of those close to us, and the welfare of foreigners lower still.

As members of a political unit, we normally agree to give each member equal weight, since there is no basis on which one person or a small group could claim to have an inherently greater entitlement than the rest and have this claim accepted. However, critics of utilitarianism such as Rawls have argued that we can justify more redistribution to the worst-off than would be implied by expected utility theory, while maintaining neutrality in an *ex ante* sense.

It turns out that the implementation of this idea in a formal sense is identical to that of rank-dependent utility under uncertainty and was first put forward around the same time, by Weymark (1981). As Ebert (…) observes, using rank-dependent weightings allows for a reconciliation between Rawls and Bentham.

Inherent discounting

One of the longest-running controversies in welfare economics has concerned the appropriateness of applying different weights to people in different generations, and, more generally of discounting future utility whoever receives it.[2] Ramsey (…), whose work is the starting point for formal analysis of intertemporal choices, rejected inherent discounting as ethically unjustified, and this viewpoint is shared by most philosophical advocates of utilitarianism. On the other hand, a good deal of evidence suggests that individuals tend to discount their own future consumption.

Before discussing inherent discounting, it is worth observing that standard expected utility suggests one reason for discounting future consumption; namely the possibility that we will not be around to enjoy it. As individuals, we face a typical annual mortality risk of around 1 per cent[3], and it makes sense to discount future utility by this amount. But at least some of the time people (most notably teenagers) discount the future much more than this.

For society as a whole, there is a comparable risk arising from the possibility of nuclear annihilation, a killer meteor and so on. The risk need not involve a total extinction of the species; it is sufficient that the disaster be great enough that 'all bets are off' in terms of calculations about the future.

With this point addressed, there remains the question of whether we do and should, discount future utility. The evidence on individual behavior is far from clear. On the one hand, there is a lot of evidence to support the idea of 'hyperbolic discounting' … However, this is offset by a notion of 'mental accounts'. Individuals may allocate resources between activities and follow inconsistent rules in different activities. For example, the same person may allocate money to an automatic saving scheme offering low or even negative real returns, while displaying hyperbolic discounting with respect to the remaining cash flow.

Leaving such phenomena to one side, the evidence for high inherent rates of discount is not strong. The most obvious market measure to use in assessing

[2] The similarity between this idea and probability weighting is discussed by Quiggin & Horowitz

[3] In the environment in which we evolved, the rate would have been higher. And an ev psych analysis suggests that what matters most is the probability of the end of reproductive life, which is higher still.

intertemporal tradeoffs is the real rate of interest on low-risk bonds (government or AAA corporate). This rate has generally been between one and two per cent and is currently around two per cent. Given that the rate of growth of average consumption per person is between one and two per cent, this is consistent with zero discounting and eta=1.

Even if individuals do display inherent discounting, that does not necessarily mean that this is appropriate as a basis for social decisions. Future individuals presumably will not share the view that utility in our time is inherently more valuable than utility in ours. In fact, as individuals, introspection and casual observation suggests that we generally regret decisions made in the past on the basis of inherent discounting. Such decisions represent selfishness on the part of our past selves at the expense of our current selves, analogous to individual selfishness with respect to others.

Implications for parametric choices

Broadly speaking, high weighting on low-end outcomes and high values of eta are substitutes. That is, the higher the weighting on low-end outcomes, the lower the value of eta required to match given observations or intuitions. Consider the income redistribution example, and suppose that the poor are given a weight twice that of the rich. Then if the redistributive project in the example, taking 1 per cent of income from the rich, and raising the income of the poor by 1 per cent, is just marginally acceptable, the required value of eta is not 1, as before, but 0.5.

Similarly, a high rate of inherent time discounting implies that to match any observed pattern of market rates, a lower value of eta is required.

Stern and the critics

The analysis in the Stern Review follows the general approach set out above. The value of eta is set to 1, which is, as noted above, the most common single choice for this parameter. The value of delta is set to 0.1 per cent, reflecting a rejection of inherent discounting, except insofar as it reflects the possibility of extinction. Similarly, there are no interpersonal weights, but changes in the income of low-income individuals and countries are weighted more highly at the margin because of the assumed diminishing marginal utility of consumption.

In Stern's analysis g is derived from the economic scenarios. Typical values of g are between 1.5 per cent and 2 per cent, so the corresponding values of r are between 1.6 per cent and 2.1 per cent.

The effect of choosing eta = 1 and delta near 0 is that concern for future generations extends more or less indefinitely into the future, when changes in welfare are expressed in terms of percentages of income. On the other hand, the discounted value of payments expressed in monetary terms declines quite fast. At a rate of 2.1 per cent, a dollar of (constant price) income received in 2100 is worth approximately 12 cents today. A income stream of a dollar a year, received for a million years into the future is worth a little under $50.

The controversy over Stern's approach has raised many concerns not all of which have been expressed clearly.. It is hard to disentangle all of these concerns, but I will try to deal with them in the following order.

(i) erroneous criticisms based on misunderstandings of the discounting procedure

(ii) criticism of the choice of delta;

(iii) criticism of the choice of eta

(iv) claims of inconsistency with observed market rates

(v) claims of inconsistency with general practice

(vi) claims of internal inconsistency

(vii) criticism of the time horizon used in discounting

(vii)

Erroneous criticisms

One part of the debate over delta can be dismissed pretty easily. Many of the critics on this point have confused delta and r, apparently assuming that delta is a discount rate, rather than a subsidiary factor determining the discount rate.

Examples of this confusion include Kling, Lomborg, and McArdle.

Criticisms of delta

Among the more serious critics, both Nordhaus and Yohe focus on the sensitivity of the results to changes in the value of delta, but do not give any specific argument for inherent discounting. Yohe does not present any argument for a high value of delta, simply observing that others have used high rates.

Nordhaus and Boyer (and hence also Nordhaus 2006) appear to backed out delta as a residual. For reasons of technical tractability, Nordhaus and Boyer want to set eta =1. On the other hand, they want r to be at least 4 per cent to match their interpretation or market data. The only way to do this is to choose high values of delta. Hence, it is probably more correct to classify Nordhaus' criticism as being based on deviation from observed market rates rather than specifically on the choice of delta.

Nevertheless, it may be worthwhile restating DeLong's response to suggestions for a large value of delta (numerical values changed)

> A δ of 2% per year is unconscionable--it means that somebody born in 1960 "counts" for twice as much as somebody born in 1995, who in turn "counts" for twice as much as somebody born in 2020; somebody born in 1960 "counts" for 256 times as much as somebody born in 2160. That's not utilitarianism

Nor is it sensible in terms of individual decisionmaking. For someone facing a zero real interest rate for savings (not an unreasonable assumption in many cases), the combination delta=2 per cent, eta =1 implies a consumption path that declines at about 2 per cent per year. To adapt Brad's example, a person beginning such a plan at age 30 would plan to halve their consumption by age 65, and halve it again by age 90. Even allowing for the caution above about EU models and actual choices, this makes no sense.

Yet Nordhaus and Boyer propose an even higher rate of 3 per cent, which is tantamount to saying that the future (certainly anyone more than two generations away from us) can go to hell for all we care, since the welfare of our great-grandchildren has about a tenth the weight we accord the current generation. Not surprisingly, this translates into a 'do nothing now' approach to global warming.

In the absence of any convincing justification for inherent discounting, the case for a low rate such as that chosen by Stern seems overwhelming. Hence, if there is a problem with the ultimate outcome it is necessary to look elsewhere in the analysis. From here on, the value of delta = 0.1 will be assumed, and discussion of the implications of other choices is conditional on this.

The intertemporal elasticity of substitution

A more plausible criticism concerns the choice of eta, the intertemporal elasticity of substitution. With these preliminaries out of the way, we can look at some of the problems that have been made of Stern's choice of eta = 1, and suggestions that higher values should have been considered.

The most direct criticism is that, in a growing economy, a low value of eta underweights the welfare of the current generation, at the expense of succeeding generations who will be much richer. This point is made most directly by Dasgupta, who considers the case when society has available an unlimited supply of projects yielding a riskless rate of return of 4 per cent. As Dasgupta shows, with eta = 1, the implied policy recommendation is that the vast majority of current income (around 97.5 per cent) should be saved in order to allow for greater consumption in the future.

The underlying problem is observed by DeLong. Looking at current savings rates and rates of economic growth, Dasgupta's estimated rate of return to marginal investment of 4 per cent seems conservative for a classical growth model based on factor accumulation. To achieve 1 per cent growth in income per person in such a model, it would be necessary to generate net additions to the capital stock equal to 25 per cent of total income each year (since 4 per cent of 25 per cent is 1 per cent). So the fact that we see more rapid growth with lower rates of net saving seems to imply that there must exist many projects with rates of return greater than or equal to 4 per cent.

However, once technical progress, generated either exogenously or through the existence of increasing returns to scale in knowledge, is taken into account, the picture changes radically. In an economy where most growth in consumption arises from technical progress, the optimal rate of saving is far lower than that derived by Dasgupta.

A more direct way of refuting Dasgupta's argument is to observe that the major premise must be false. If there existed an infinite supply of projects with riskless returns of 4 per cent, the rate of return on riskless bonds would have to equal 4 per cent, rather than the 1 to 2 per cent observed in practice. Although this difference may appear small, it is critical in practice

(iv) claims of inconsistency with observed market rates

A number of critics, notably including Nordhaus (..) have argued that the discount rate implied by Stern's procedure is inconsistent with observed market data. As has been observed several times above, this claim is incorrect with respect to the most obviously relevant data point, the real rate of return on low-risk bonds.

On the other hand, application of the expected utility model with parameters like those of Stern produces a severe underestimate of the expected rate of return to equity investments, and this is Nordhaus' main concern. In effect, this is a restatement of the equity premium puzzle, along with an implicit presumption that the market rate of return to equity, and not the market rate of return to debt, is the 'correct' measure of the social rate of time preference.

Although defensible, this seems implausible. Most resolutions of the equity premium puzzle imply that the puzzle mainly reflects an excessive risk premium for equity rather than an inadequate rate of time preference.

However, the problems observed here reflect the underlying difficulties of the expected utility model. No single choice of eta and delta produces results consistent with actual observations on choices over time and risk, and intuitions about redistributions between individuals. One manifestation of this is the equity premium puzzle. Given plausible estimates of risk attitudes, either the real bond rate is much too low (it's typically between 1 and 2 per cent) or the real rate of return to equity is much too high (it's typically 6 to 8 per cent). The model suggests that the two should differ by no more than half a percentage point.

Logically, any proposition can be deduced from a contradiction. So it is easy, in the present context, to combine plausible empirical propositions with standard EU analysis and derive absurd results. In general, this is not a useful way to proceed.

(v) claims of inconsistency with general practice in benefit-cost analysis

Tol makes the observation that British government generally uses discount rates for benefit-cost analysis that are much higher than those proposed by Stern. This is a neat debating point, but has little practical relevance. The political structure of project appraisal is that estimates of costs and cash flows rely on inputs from project proponents that are almost always over-optimistic. Treasury controls the choice of discount rate and uses it to adjust for downside risk as well as for discounting. Hence,the official discount rate is substantially higher than the true social rate of discount.

(vi) claims of internal inconsistency

Tol also observes that whereas Stern uses eta =1 for discounting, estimates of the cost of disaster are derived from Nordhaus and Boyer who use eta = 4 for this purposes. It appears that this inconsistency did not originate with Stern, but with Nordhaus and Boyer. Nevertheless, it would seem appropriate to address it.

(vii) criticism of the time horizon used in discounting

The most plausible criticism of low rates of discount is that they require us to take account of developments more than 100 years into the future about which we can in practice, know very little. This is a reasonable criticism, but its main effect is to point up the limitations of utilitarian benefit-cost analysis for a problem like global warming.

We know that the effects of global warming will be felt far into the future. We can either mitigate these effects, at very modest cost to ourselves, or leave the problem future generations, whose technological capacities are unknown, but presumably greater than our own. Perhaps our descendants will be able, at very low cost, to resuscitate species we have driven to extinction and restore ecosystems we have destroyed. Perhaps not. There is no easy way of getting useful probability and cost nubmers here.

One partial solution might be to end the analysis at, say 2050 or 2100, with future effects being measured as a diminution in the capital stock (including natural capital). Although logically equivalent to the discounting procedure employed by Stern and his critics, this might turn out to be more tractable and intuitively comprehensible.

Concluding comments

Criticism of Stern has focused on the claim that the parameters used in discounting are extremely low, yielding implausible results. In fact, the choice of eta =1 is standard, and both lower and higher values are commonly considered in sensitivity analysis. Stern's choice of delta = 0.1 per cent is primarily the result of applying the standard utilitarian view that all people count equally. If this view is accepted, the pure rate of time discount, reflecting the probability of social extinction, must be close to zero, and there is nothing remarkable about the parametric value delta = 0.1 per cent.

The real difficulty here is that we are pushing economic analysis to its limits, in an area where fundamental problems, such as the equity premium puzzle remain unresolved. Economists can help to define the issues, but it is unlikely that economics can provide a final answer.