Geir B. Asheim *

# Ethical preferences in the presence of resource constraints

The interest in the concept of a *sustainable development* arises from a concern that natural and environmental resources may not be managed in a manner that secures the well-being of future generations. By letting each generation's capital and resource management be made according to its *subjective preferences* and by assuming that each generation through these subjective preferences exhibits altruism for its descendants, then this issue translates into a concern that the altruism may not be sufficient to secure the well-being of future generations.

An example where these concerns are well-founded is constructed by letting the subjective preferences of each generation $t$ be represented by

$$w_t = u_t + bw_{t+1} = \sum_{s=t}^{\infty} b^{s-t}u_s, \quad 0 < b < 1.$$

Here the *subjective welfare* of generation $t$ ($w_t$) depends on its own *utility* ($u_t$) and the subjective welfare of its immediate descendents. The utility of generation $t$ is assumed to be an unquestionable indicator of the well-being that is generated from the situation in which it lives. Due to its recursive form these subjective preferences are equivalent to *discounted utilitarianism*. The fact that the utility discount factor ($b$) is smaller than 1 means that future utility is discounted. There is a two-fold justification for these subjective preferences: (1) They are consistent with the observation that parents care for the welfare of their children. (2) They are consistent with an objective function that is very commonly applied and for which Koopmans (1960) has provided an axiomatic foundation.

It is not hard to construct economic models where discounted utilitarianism leads to ethically unacceptable results due to the presence of resource constraints. The most well-known such example is the Dasgupta-Heal-Solow (D-H-S) model of capital accumulation and resource depletion (Dasgupta & Heal (1974; 1979, Ch. 10) and Solow (1974)). In the simple version that will be considered here, it is a model with non-overlapping generations of constant size and no technological progress. A stock of reproducible capital and a flow of resource extraction enters a production function. Production is split into consumption ($c_t$) and capital accumulation.

* University of Oslo. I thank Ottar Mæstad and Haakon Vennemo for valuable comments.

*Figure 1*

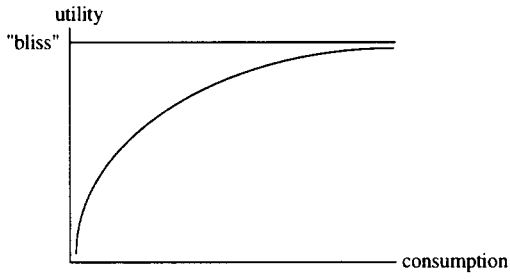



*Figure 2a*

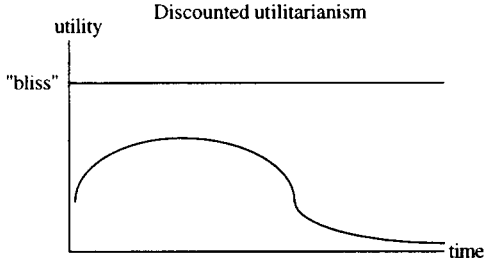
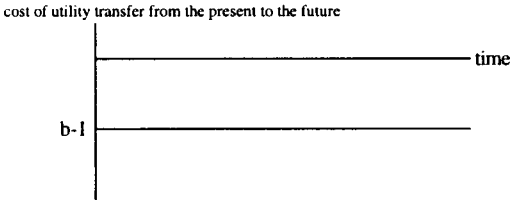



*Figure 2b*




No production is feasible without resource extraction, while total resource extraction is constrained by the availability of a finite stock since the resource is assumed to be non-renewable. However, with positive initial stocks of capital and resource, positive and constant consumption is feasible.[1] This is achieved by having an increased stock of reproducible capital make up for a diminishing flow of resource extraction. Let utility depend on consumption: $u_t = u(c_t)$. I will assume that $u(\bullet)$ is strictly increasing, strictly concave and bounded above, as illustrated by Figure 1.[2] The upper bound on utility values will be called 'bliss'. Under these assumptions, if the subjective preferences are represented by $w_t = u_t + bw_{t+1}$, then consumption is forced to approach zero as time goes to infinity. This manifestly unjust outcome is obtained no matter how small the discounting is, i.e., no matter how close the discount factor $b$ is to 1. Any constant discounting of the future leads to the same qualitative result. The optimal utility path according to discounted utilitarianism is illustrated in Figure 2a.

**Ethical preferences**

In order to achieve an ethically acceptable outcome in the D-H-S model, *ethical preferences* need to be introduced. The terminology, 'ethical' vs 'subjective' preferences is taken from Harsanyi (1955). While the term 'subjective preferences' is here meant to capture 'selfish' altruism – in the sense that a generation is motivated to contribute to the welfare of its children because it leads to increased welfare of the contributor – the term 'ethical preferences' is meant to capture preferences which are costly to abide by and which, as a consequence, will not be accepted unless they are imposed as a moral obligation.

Harsanyi (1953) and Rawls (1971)[3] argue that ethical preferences on intergenerational

1. This holds in continuous time if the stock of capital and the flow of resource extraction enter into a Cobb-Douglas aggregate production function where the factor elasticity of the capital stock exceeds the factor elasticity of the resource flow. In the present discussion of the D-H-S model, I assume that the model has the same properties as in continuous time with such a Cobb-Douglas function. For an analysis of the D-H-S model in discrete time, see Dasgupta & Mitra (1983). For a general analysis of the feasibility of positive and constant consumption despite a resource constraint, see Cass & Mitra (1991).
2. The assumption that $u(\bullet)$ is bounded above is needed for the Chichilnisky and Ramsey criteria considered later.
3. Due to the influence of Rawls' (1971) *A Theory of Justice*, also on questions of intergenerational justice, I will relate much of my discussion to his book.

distributions should stem from principles of justice that generations would agree on in the original position. Moreover, Rawls (1971: 4) supports the view that ethical preferences should be evaluated by their implications in specific economic environments by posing the question: Are the distributional consequences acceptable given our moral intuitions?[4] This is in line with his concept of a *reflective equilibrium.*

Based on Rawls' (1971) view, I will here evaluate different forms of ethical preferences by showing their implications in the D-H-S model. That a specific form of ethical preferences yields ethically acceptable implications in the D-H-S model does not, of course, mean that it is universally applicable. However, since the D-H-S model is very often used to show that discounted utilitarianism produces ethically unacceptable outcomes, it seems natural to subject alternative criteria to the same test.

Before undertaking this exercise, it is useful to present a taxonomy which can be applied for classifying the different forms of ethical preferences that will be considered. On the one hand, ethical preferences may differ with respect to the object on which they are defined:

- The ethical preferences may be defined on the flow of subjective welfares: $(w_1, w_2, w_3, \ldots)$. This is in accordance with what Sen (1977) calls *welfarism*: «The general approach of making no use of any information about the social states other than that of personal welfares generated in them ... .» I will assume that $w_t$ is interpersonally level comparable.
- The ethical preferences may be defined on the flow of utilities: $(u_1, u_2, u_3, \ldots)$. I will assume that $u_t$ is interpersonally level and unit comparable.

On the other hand, ethical preferences may differ with respect to what kind of ordering they yield:

- The ethical preferences may yield a complete preordering, implying that if an intergenerational distribution is *maximal* [5] with respect to the preordering, then it is weakly preferred to all feasible distributions.
- The ethical preferences may yield a partial preordering,[6] allowing for the possibility that maximal intergenerational distributions may not be comparable. Intergenerational distributions that are not maximal with respect to the preordering are deemed ethically unacceptable.

It must be required that the ethical preferences do not depend on the stocks of capital and resource. This is especially important in a model (like the D-H-S model) where positive consumption entails depletion of a non-renewable resource. Because otherwise, any feasible flow of utilities (or welfares) with positive consumption at any time can be realized by conditioning in a particular way the capital and resource management on the remaining stock of the resource.[7] Moreover, it must be

4. This approach is in principle supported also by e.g. Koopmans (1967), Mishan (1977), and Dasgupta & Heal (1979, pp. 308–311).
5. A distribution is *maximal* if there is no alternative distribution that is strictly preferred to it.
6. A partial preordering is reflexive and transitive, but not complete.
7. This restriction seems to preclude that the ethical preferences at time $t$ are endogenous in the sense of being dependent on the history up to time $t$. Note, however, that for each of the Ramsey, Solow, Dalton, and Calvo criteria considered later, it is possible to express each generation's optimal capital and resource management as a function of its inherited stocks of capital and resource. This can be done even though the ethical preferences do not depend on stocks.

required that generations have identical ethical preferences. Because otherwise, any feasible flow of utilities (or welfares) can be realized by letting the preferences change. This means that the ethical preferences must not depend on *absolute* time. Finally, generations must perform consistent planning in the sense of taking into account that the same ethical preferences will be applied by future generations.

**The Chichilnisky criterion**

Chichilnisky (1996) has suggested a class of ethical preferences that are based on the axioms that neither 'the present' nor 'the future' should be dictated. Ethical preferences in this class are defined on $(u_1, u_2, u_3, \dots)$ and depend both on the flow of utilities over time and the long-run behavior of utility values. Following Beltratti, Chichilnisky & Heal (1995), I here consider the following representation (where the discount factor $\delta$ need not coincide with the discount factor $b$ of the subjective preferences):[8]

$$\theta \sum_{t=1}^{\infty} \delta^{t-1} u_t + (1-\theta) \liminf_{T\to\infty} \inf_{t\geq T} u_t, \quad 0<\theta<1, \quad 0<\delta<1.$$

Hence, the ethical preferences are represented by a weighted sum of the sum of discounted utilities and the level of utilities that is approached when time goes to infinity.

Beltratti, Chichilnisky & Heal (1995, p. 149) «... feel that this formulation is in tune with the concerns of those who write about sustainability and our responsibilities to future generations.» However, any of the future generations that such writers have concern for will live in *finite* time. Hence, the Chichilnisky criterion is in tune with these concerns only if the weight on the limiting properties of the utility flow is instrumental in securing the well-being of future generations. Therefore, it seems interesting to pose the question: What are the implications of imposing the Chichilnisky criterion as ethical preferences in the D-H-S model? Note that Alistair Ulph in his discussion of Beltratti, Chichilnisky & Heal (1995) argues that this is an important exercise: «I would like to see the implications of the proposed new welfare criterion worked out in a wider class of models, and in particular models of the type ... involving non-renewable resources and for which we know that the discounted utilitarian criterion with high utility discount rates gives unpalatable outcomes» (Ulph (1995, p. 171).

In the D-H-S model, 'bliss' can be approached in the limit as time goes to infinity as long as there are positive stocks of capital and resource at any finite time. Hence, the second term of the objective function representing the Chichilnisky criterion is not precluded from attaining its maximum value as long as the stocks are positive. If both stocks are not positive, consumption must converge to zero as time goes to infinity. This means that the second term of the objective function works as an instrument for keeping stocks positive. Now, the program that maximizes the sum of discounted utilities – i.e., the first term of the objective function representing the Chichilnisky criterion – implies that stocks are positive at any finite time. Hence, 'bliss' can be approached as time goes to infinity even if the concern for future generations is postponed. Therefore, such a postponement is desirable according to the Chichilnisky criterion since it does not affect the second term of the objective function, while it allows for an increase of the first term. On the other hand, ethi-

8. If $u_t$ is a convergent sequence, then lim inf $u_t$ = lim $u_t$. Note that lim inf $u_t$ is defined even if $u_t$ is not convergent.

cal welfare is decreased if the concern for future generations is postponed indefinitely since then 'bliss' is not approached as time goes to infinity. The conclusion is that *no optimal distribution exists.*[9]

It is important to note that this non-existence problem does not arise due to technical issues of a mathematical nature. It arises because the limiting properties of the utility flow as time goes to infinity cannot be used as a proxy for the well-being of future generations that live in finite time. In fact, in the D-H-S model a utility path that is 'nearly' optimal according to the Chichilnisky criterion will be close to the utility path that is optimal according to discounted utilitarianism.[10] Beltratti, Chichilnisky & Heal's (1995, p. 150) claim – that their «approach is fully consistent with earlier discussions of sustainability» – is therefore puzzling.

**The Ramsey criterion – Classical utilitarianism**

Ramsey (1928) suggests to apply classical utilitarianism to intertemporal distributions. This means that the ethical preferences are defined on $(u_1, u_2, u_3, \ldots)$ and are represented by

$$\sum_{t=1}^{\infty} (u_t - u),$$

where $u$ is the 'bliss' utility level. Dasgupta & Heal (1979, Ch. 10) consider the Ramsey criterion and show that an optimal utility path exists under given assumptions on the utility function. When it exists, the optimal utility path is always increasing and approaches 'bliss' as time goes to infinity. These properties of the optimal utility path – which are caused by the fact that investment in capital is always productive in the D-H-S model in the sense of yielding a positive net return – are illustrated in Figure 3a.

It will prove useful to describe criteria in terms of their relation to the cost of intergenerational transfers. Therefore, say that a transfer of utility from one generation to another has *negative cost* if the provider loses less utility than the receiver gains. Due to positive productivity of capital investments, a given sacrifice of *consumption* can be turned into a greater gain of consumption for the next generation. Hence, along a path with constant consumption (so that total and marginal utility is constant), a given sacrifice of *utility* can be turned into a greater gain of utility for the next generation. In this case, there is a negative cost of utility transfer from one generation to the next. Note that the cost of utility transfer is related to capital productivity, growth in consumption, and the concavity of the utility function; not to incentive constraints as in the case of intragenerational redistribution.

A utility path that is optimal according to discounted utilitarianism has a negative cost of utility transfer from one generation to the next since a loss of $b$ (<1) for the provider

9. Existence may be ensured if $\delta$ is time variant and the utility discount *rate* $((-d\delta/dt)/\delta)$ goes sufficiently fast to zero as time goes to infinity. However, in such a case the second term of the objective function is not needed since maximization of the first term ensures that 'bliss' is approached as time goes to infinity. Furthermore, given our requirements, $\delta$ can only depend on *relative* time, not *absolute* time. When $\delta$ is not time invariant, this in turn creates a time inconsistency problem in the sense that some generation $t > 1$ will not want to continue the optimal path at time 1. Whether existence can be restored when generations plan consistently appears to be an open question.

10. The following is precise statement of this result. The supremum of the Chichilnisky objective function is the weighted sum of the value of the path optimal according to discounted utilitarianism and the value of a path approaching 'bliss' as time goes to infinity. Consider for any $n$ a path with a value according to the Chichilnisky objective function less than $1/n$ from the supremum. As $n$ goes to infinity, this sequence of paths approaches in the topology of pointwise convergence the path that is optimal according to discounted utilitarianism.