

## **Discounting and the Time Preference Rate**

### **Author**

Creedy, John, Guest, Ross

### **Published**

2008

### **Journal Title**

Economic Record

### **DOI**

[10.1111/j.1475-4932.2008.00450.x](https://doi.org/10.1111/j.1475-4932.2008.00450.x)

### **Rights statement**

© 2008 Blackwell Publishing. This is the author-manuscript version of the paper. Reproduced in accordance with the copyright policy of the publisher. The definitive version is available at [www.interscience.wiley.com](http://www.interscience.wiley.com)

### **Downloaded from**

<http://hdl.handle.net/10072/21590>

### **Link to published version**

<http://www.wiley.com/bw/journal.asp?ref=0013-0249>

### **Griffith Research Online**

<https://research-repository.griffith.edu.au>

# Discounting and the Time Preference Rate\*

John Creedy and Ross Guest

The University of Melbourne and Griffith University

## Abstract

This paper provides an analytical review of the evaluation of alternative time streams of consumption and the closely related concept of time preference. The potential sensitivity of comparisons, especially to the choice of time preference rate and elasticity of marginal valuation, is demonstrated. The nature of time preference, based on an axiomatic approach, is then discussed. The analysis of optimisation over time leads to the concept of the social time preference rate, and a difficulty with using this rate is highlighted. Approaches giving rise to declining discount rates over time are discussed, including alternative welfare functions and the role of uncertainty. This is followed by a critique of methods used to ‘estimate’ a time preference rate. Finally, complications introduced by non-income differences between individuals are examined. Emphasis is placed on the central role of value judgements.

---

\*We are grateful to Denis O’Brien, Tony Scott and two referees for comments on an earlier version of this paper.

# 1 Introduction

The aim of the present paper is to review a central issue in the evaluation of alternative time streams – that of discounting. The context is one in which evaluation of a public project is made by a disinterested judge, that is someone who has no personal interest in the outcomes. Investment projects typically involve a present cost incurred in order to achieve future benefits.<sup>1</sup> These might be in the context of investments in health technology, civil engineering projects, or environmental protection. It is therefore necessary to evaluate alternative outcomes, involving different time streams of net benefits. Despite the long-standing nature of the problem, it remains controversial and even the basic issues are far from being settled.<sup>2</sup> One of the problems concerns a lack of clarity over the concepts. Another difficulty arises from the fact that there is no escape from fundamental value judgements, while protagonists on different sides of debates often conceal their value judgements. Hence it is important to be clear about precisely how they enter the calculations and how they may be specified. Emphasis is placed on the analytical issues rather than providing a comprehensive literature review.

Section 2 sets the scene by considering social evaluations based on the concept of a social welfare function: this is dominant in the literature concerned with evaluating public projects. This form of welfare function involves, as well as attaching different weights to different levels of consumption irrespective of their timing, the discounting of future flows using what is called a ‘pure time preference rate’. There are alternative views about the way to proceed. One approach is simply to say that the social welfare function is meant to represent alternative value judgements and therefore results should be reported for alternative time preference rates. Some economists attempt to impose their own value judgements, using rhetorical arguments suggesting for example that pure time preference is in some sense ‘ethically indefensi-

---

<sup>1</sup>In any exercise of this kind there are obviously huge problems associated with measurement issues and uncertainty about the future. But these are not the focus of attention here.

<sup>2</sup>For example, the controversial nature of discounting is demonstrated by the debate over the Stern Report (2006) on climate change; see, for example, Carter *et al.* (2006), Dasgupta (2006), Nordhaus (2006) and Varian (2006).

ble'. Thus Ramsey (1928, p.543) stated that discounting utility over time is 'a practice which is ethically indefensible and arises merely from the weakness of the imagination'. Pigou (1932, p. 25) also argued that time preference implies that 'our telescopic faculty is defective'. More recently an objection to discounting has been made in terms of overlapping generations, the criticism being that it is 'unethical' to impose the preferences of the current generation on the utility of people who are not yet born (Padilla, 2002).<sup>3</sup>

However, it is desirable to have a clear understanding not only of what is implied by pure time preference – or its absence – but what specific value judgements may lie behind it. That is, it is useful to appreciate how time preference can arise from more basic axioms stating value judgements in a clear way. Section 3 discusses an axiomatic approach to time preference, based on the argument of Koopmans (1960).

Section 4 turns from social evaluations of exogenous time profiles to decisions regarding the socially optimal allocation of resources over time. It therefore concerns the planning, again by an independent judge, of optimal saving and consumption patterns. However, it uses the same kind of social welfare function. Section 5 returns to the evaluation of alternative streams in the context of cost-benefit analyses. It discusses the concept of the social time preference rate and highlights a problem with its application. Situations giving rise to the use of decreasing discount rates over time are examined in section 6, including the use of an alternative formulation of the social welfare function involving 'sustainable preferences' and the effect of uncertainty. Section 7 introduces some complications arising from attempts to allow for the fact that income units differ in size and composition. Some approaches to measurement are examined in Section 8. The interpretation of orders of magnitude regarding one aspect of value judgements is discussed in Section 9. Brief conclusions are in Section 10.

---

<sup>3</sup>It is a small step to the argument that it is unethical to impose a constant discount rate over the lifetime of an individual because an individual's life can be divided into the current self and the future self who are effectively two different persons; see Caplin and Leahy (2000).

## 2 Social Evaluations

Suppose it is required to evaluate a time stream  $C = [c_1, c_2, \dots]$  of consumption. For simplicity, it is assumed that the population consists only of individuals (rather than families), that the size of the population remains unchanged over time, and that consumption is the only economic variable considered to be relevant by the judge. These assumptions are relaxed in section 6.1 below. The term  $c_t$  refers to aggregate consumption in period  $t$ . Hence there is, by assumption, no concern for within-period inequality among individuals. An evaluation cannot avoid the use of value judgements. Hence, the usual approach is to examine the implications of adopting a range of value judgements.

### 2.1 A Social Welfare Function

Consider social evaluations based on additive Paretian social welfare functions, that is functions for which social welfare,  $W(C)$ , is a weighted sum of consumption values in each period and an increase in consumption in one period, without a reduction in any other period, is regarded as an improvement. There are two components to such functions, involving first a weight,  $U(c_t)$ , to be attached to each  $c_t$ , irrespective of the time period,  $t$ , and second a view about the timing of consumption, reflected in a constant rate of pure time preference,  $\rho$ .<sup>4</sup>

Social welfare functions of this class take the form:<sup>5</sup>

$$W(C) = \sum_{t=1}^T U(c_t) \left( \frac{1}{1+\rho} \right)^{t-1} \quad (1)$$

The effects of discounting alone can be seen in Figure 1, which plots the discount factor,  $\left( \frac{1}{1+\rho} \right)^{t-1}$ , against  $t$ , for several alternative values of  $\rho$ : in other words it shows how the present value of \$1 falls as the time period increases. This rate of decline is clearly highly sensitive to the choice of  $\rho$ .

---

<sup>4</sup>The use of falling discount rates with time is considered below.

<sup>5</sup>The presentation here uses a discrete time framework, but conversion to continuous time would not affect the results.

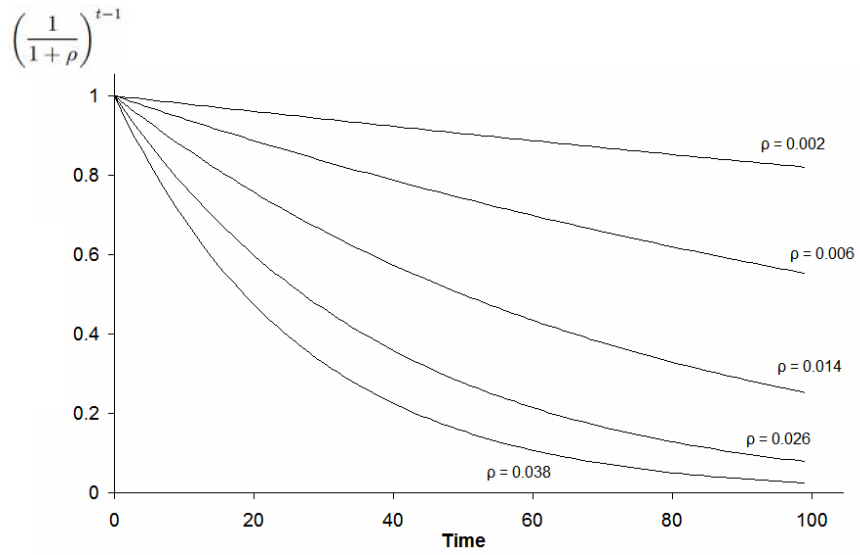


Figure 1: The Discount Factor and Time Period

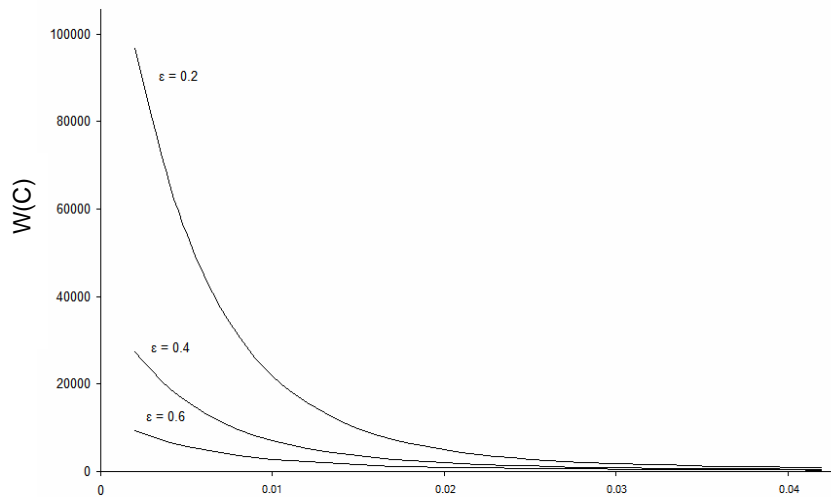


Figure 2: Sensitivity of  $W(C)$  to Choice of Epsilon (Less Than 1):

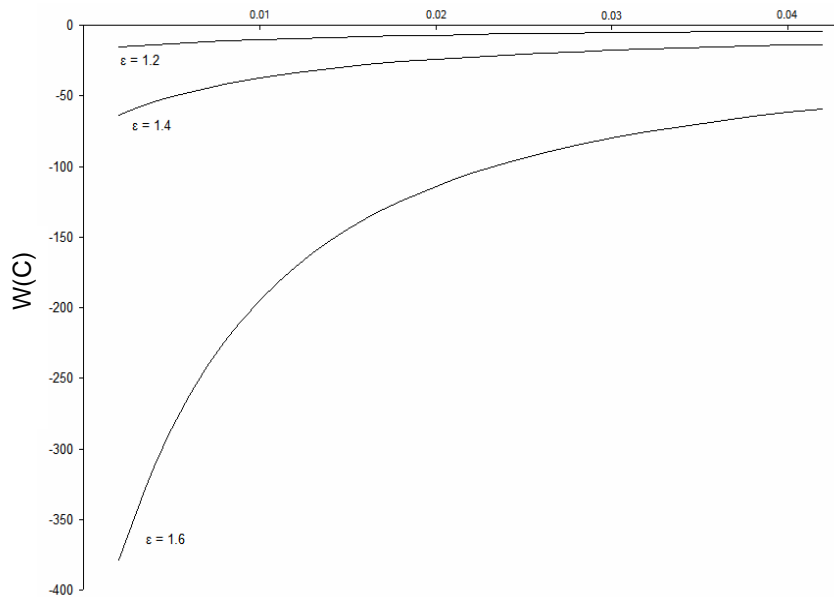


Figure 3: Sensitivity of  $W(C)$  to Choice of Epsilon (Greater Than 1):

The weighting function  $U$  – a cardinal measure of the contribution to  $W$ , before discounting, of period  $t$ 's consumption – is sometimes called a utility function. Hence the pure time preference rate is sometimes also called a ‘utility discount rate’. However, this terminology is somewhat misleading unless it refers to a single-person framework. The social welfare function  $W$  does not represent the wellbeing of society: above all, it does not represent ‘society’s views’, but those of an independent judge.<sup>6</sup>

It is usual to consider the implications of types of  $U(c)$  which reflect a decreasing marginal valuation,  $dU(c)/dc$ . The rate at which the marginal valuation falls as  $c$  increases is described by the ‘elasticity of marginal valuation’,  $d \log(dU(c)/dc)/d \log c_t$ . This elasticity is negative for concave  $U$ . This value judgement therefore reflects adherence to a multi-period ‘principle of transfers’ whereby (in the absence of discounting) a transfer of consumption from high to low consumption periods, so long as their relative ranks do not change, is considered an improvement.

---

<sup>6</sup>Indeed, it is known from Arrow’s Impossibility Theorem that there is no consistent way of aggregating preferences, such that certain axioms are satisfied.

Consideration of alternative value judgements regarding  $U$  is simplified by the use of constant-elastic functions. This means that  $U$  takes the form:

$$U(c_t) = \frac{c_t^{1-\varepsilon}}{1-\varepsilon} \quad (2)$$

where the term  $\varepsilon$  is the absolute value of the elasticity of marginal valuation, since  $dU(c)/dc = c_t^{-\varepsilon}$  and  $\log(dU(c)/dc) = -\varepsilon \log c_t$ . Another way of describing  $\varepsilon$  is that it measures the judge's degree of constant relative aversion to variability over time: in the absence of discounting the judge would prefer to see a smooth consumption stream.

The implications of adopting different value judgements can therefore be examined, for a given consumption stream, by calculating  $W$  for alternative values of  $\varepsilon$  and  $\rho$ . Consider a consumption stream over 250 periods, where the initial value is 30 units and there is smooth growth at the constant rate of 2.3 per cent per period. The values of  $W(C)$  are highly sensitive to the choice of  $\varepsilon$ , as shown in Figures 2 and 3, where each profile shows the variation in the present value of  $W(C)$  as  $\rho$  is increased, for a given value of  $\varepsilon$ . Figure 2 shows the reduction in the present value as  $\varepsilon$  is increased from 0.2 to 0.6, while Figure 3 shows variations for values of  $\varepsilon > 1$ , for which  $W(C)$  is negative.

The introduction of the terms  $\rho$  and  $\varepsilon$  in the social welfare function makes it clear that these reflect the value judgements of a hypothetical judge or decision maker. However, there is a literature attempting to 'estimate' values using a variety of methods. These are discussed in section 8.<sup>7</sup>

## 2.2 Comparing Alternative Time Streams

In view of the sensitivity of present values, it cannot be expected that alternative projects have the same ranking, independent of the choice of elasticity of marginal valuation and time preference rate. Consider the two profiles A and B in Figure 4, where B has the fastest constant growth rate of 1.6 per cent, compared with A of 0.9 per cent, but the starting value of B is 5 while

---

<sup>7</sup>Some authors decompose  $\rho$  to include a term depending on the perceived probability of extinction: see Pearce and Ulph (1998) and Stern (2006).



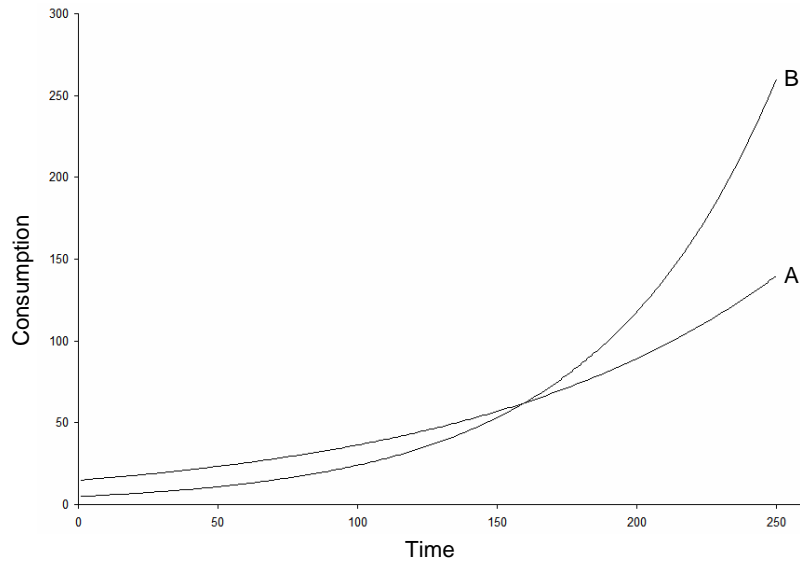


Figure 4: Two Consumption Profiles With One Intersection

that of A is 15. The profiles intersect only once. Time profile B is expected to dominate A only for relatively low values of  $\rho$ , though the particular value of  $\rho$  for which the ranking changes depends crucially on the choice of  $\varepsilon$ . Present values,  $W(C)$ , are shown in Figure 5 for  $\varepsilon = 0.6$ , showing that profile B is preferred for low values of  $\rho$ . However, it is found that for elasticity of marginal valuation, that is higher aversion to variability (values of  $\varepsilon > 0.88$ ), the flatter profile A is preferred to B for all  $\rho$  varies.

More complex comparisons may result from more variable time profiles, making the choice of alternative streams more sensitive to the choices of  $\varepsilon$  and  $\rho$ . Consider Figure 6, where time stream A results from a constant growth rate of 2.3 per cent (starting from 10 units), but profile B results from a fixed trend rate of growth of 1.8 per cent (starting from 4 units) combined with a cyclical growth component having an amplitude of 5 per cent and a wavelength of 165 periods. The consumption profiles intersect twice, so it is likely that stream A has the highest value of  $W(C)$  for both low and high values of  $\rho$ , while stream B is likely to dominate (that is have a higher  $W(C)$  value) for intermediate values, though the precise values are again likely to be sensitive to the choice of  $\varepsilon$ . An example is given in Figure 7 for  $\varepsilon = 0.2$ ,

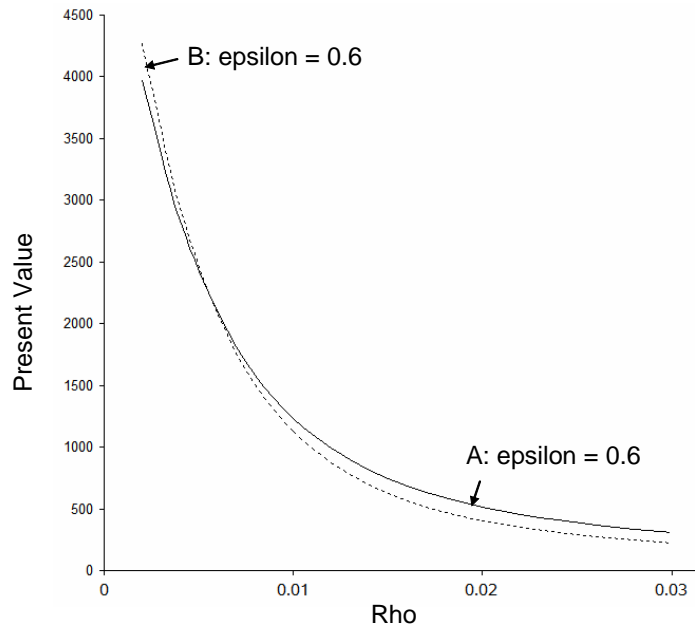


Figure 5: Present Value,  $W(C)$ , of Consumption Profiles in Figure 4 for  $\varepsilon = 0.6$  and Alternative  $\rho$

where project A is indeed preferred for high and low values of  $\rho$ .

The welfare function in (1) represents a particular set of value judgements, as well as those giving rise to pure time preference: the evaluation function is additive and Paretian. Alternative views about the desirable evaluation of a time stream of consumption are obviously possible, and professional economists cannot make prescriptions about the form to be used, but can only investigate the implications of adopting alternative forms. An alternative approach is discussed in section 6.1.

### 3 Existence of Time Preference

The social welfare function in (1) simply starts from the position that the independent judge whose value judgements are examined has positive pure time preference. However, the precise nature of the value judgements underlying time preference are not immediately obvious. The question considered here is whether time preference is implied by a clear set of axioms describing an

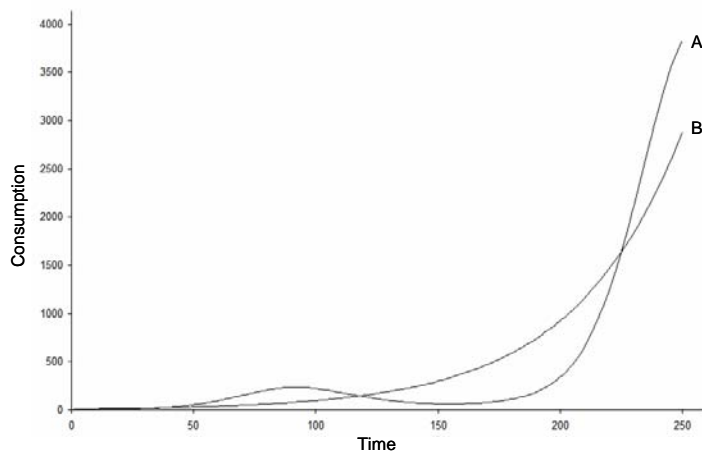


Figure 6: Two Consumption Profiles With Two Intersections

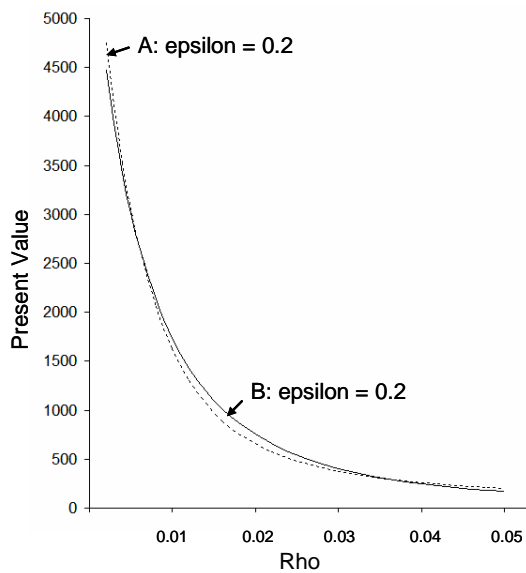


Figure 7: Present Value,  $W(C)$ , of Consumption Profiles in Figure 6 for  $\varepsilon = 0.2$  and Alternative  $\rho$

independent judge's or social planner's value judgements about time profiles of consumption. This makes it easier to identify precisely why judges may differ in their attitudes towards time preference. The following discussion is a simplified version of the argument put forward by Koopmans (1960).<sup>8</sup> This is not in any sense a 'proof' that time preference does exist or is necessary: rather, it seeks to understand the nature of the axioms required and thus to make more transparent the nature of time preference.

### 3.1 An Axiomatic Approach

Consider an independent judge with an ordinal evaluation function, given by  $P(C) = P(c_1, c_2, c_3, \dots)$  and defined over a time stream of consumption represented by the vector,  $C = [c_1, c_2, c_3, \dots]$ . It is simply assumed that this function has the usual properties of evaluation functions, such as monotonicity and transitivity. For simplicity, it is assumed that the population consists only of individuals (rather than families), that the size of the population remains unchanged over time, and that consumption is the only economic variable considered to be relevant by the judge. The term  $c_t$  refers to aggregate consumption in period  $t$ . Hence there is, by assumption, no concern for within-period inequality among individuals.

Stated informally, the *continuity* axiom states that any slight variation in  $C$  does not lead to big changes in the valuation of  $C$ , while a *boundedness* axiom states that paths  $C_A$  and  $C_B$  exist such that  $P(C_A) \leq P(C) \leq P(C_B)$ . If alternative paths were to produce unbounded values of  $P$ , they could not be ranked.<sup>9</sup>

The *sensitivity* axiom says that if paths  $C_0$  and  $C_1$  differ in only the first period, then  $P(C_0) \neq P(C_1)$ . Essentially this is stating that the first period

---

<sup>8</sup>Other demonstrations are available. Marina and Scaramozzino (2000, p.6) provided an interesting analysis of growth in an overlapping generations framework. They stated that, 'a social rate of pure time preference is justifiable on purely ethical grounds'. A clearer statement of what the authors showed is that *if* the objective of maximising average steady-state consumption per capita is adopted, then *an implication* of this *ethical value judgement*, combined with a model containing productivity and population growth, is that positive time preference exists that does not reflect myopia.

<sup>9</sup>Alternative (positive) time streams of consumption over an infinite period could not be compared in the absence of time preference, because they would be unbounded.

matters, in that it cannot be swamped by all other periods. Without the sensitivity axiom, a small gain to each of an infinitely large number of future periods, achieved at the expense of reducing consumption in the present period to zero, would be regarded as acceptable.

A *non complementarity* (or *independence*) axiom states that if two time streams differ only by the first period, their ranking does not depend on the form of the remaining stream. Here, it is convenient to introduce the notation  $C^{[2]} = (c_2, c_3, c_4, \dots)$ , so that  $C = (c_1, C^{[2]})$ . Hence, for two time profiles  $C_0 = [c_{0,1}, c_{0,2}, c_{0,3}, \dots]$  and  $C_1 = [c_{1,1}, c_{1,2}, c_{1,3}, \dots]$ , where  $c_{k,t}$  refers to consumption in the  $t$ th time period and the  $k$ th time stream, independence implies that if:

$$P(c_{0,1}, C_0^{[2]}) \geq P(c_{1,1}, C_0^{[2]}) \quad (3)$$

then:

$$P(c_{0,1}, C_1^{[2]}) \geq P(c_{1,1}, C_1^{[2]}) \quad (4)$$

and vice versa.

Finally, a *stationarity* axiom states that if paths  $C_0$  and  $C_1$  have the same consumption in the first period, so that  $c_{0,1} = c_{1,1} = c_1$ , then the ranking:

$$P(c_1, C_0^{[2]}) \geq P(c_1, C_1^{[2]}) \quad (5)$$

implies also that:

$$P(C_0^{[2]}) \geq P(C_1^{[2]}) \quad (6)$$

Hence the rankings of the alternative streams (with a common first element) remain unchanged if they are simply moved earlier one period in time.

Having stated the axioms, consider two time paths  $C_1$  and  $C_2$  such that  $c_{1,t} > c_{2,t}$  for all  $t$ , and all  $c_{k,t}$  are positive consumption levels ('all goods are good'). It must therefore be the case that  $P(C_1) \geq P(C_2)$ . Suppose there are two other time streams,  $C_3 = (c_{3,1}, C_1)$  and  $C_4 = (c_{4,1}, C_2)$  where  $c_{3,1} = c_{4,1}$ . Hence streams  $C_3$  and  $C_4$  have a common first period's consumption level, and thereafter have precisely the same streams, respectively, as  $C_1$  and  $C_2$ . The stationarity axiom therefore implies that  $P(C_3) \geq P(C_4)$ .

By definition, the paths  $C_3$  and  $C_4$ , by having a common first element, are less different than  $C_1$  and  $C_2$ . Since, from above, each period matters,

this implies that:

$$P(C_1 - C_2) > P(C_3 - C_4) \quad (7)$$

This property implies that the difference is smaller, the more distant in time it is: this is referred to as ‘time perspective’; see Koopmans, Diamond and Williamson (1964).

Next, consider alternative streams such that  $C_1$  and  $C_2$  differ only in the first time period, such that  $c_{1,1} - c_{2,1} = 1$ . Hence the streams  $C_3$  and  $C_4$  differ only in their second period, by the same amount. Using (7) it can be seen that:

$$P(1, 0, 0, 0, 0, \dots) > P(0, 1, 0, 0, 0, \dots) \quad (8)$$

Hence, with only one unit of consumption available, there is a preference for having this in the first period, rather than having nothing in the first period and waiting to consume the unit in the second period. There is therefore a preference for bringing the consumption forward from the second to the first period. This result clearly implies pure time preference.

### 3.2 A Measure of Pure Time Preference

It is necessary to have a measure of the extent of this pure time preference. Consider for simplicity the two-period case. Time preference can be interpreted in a diagram with period 2’s consumption on the vertical axis and period 1’s consumption on the horizontal axis, using the concept of social indifference curves, along which  $P$  is constant. In general, the absolute slope of the social indifference curve, the marginal rate of substitution of period 1’s consumption for period 2’s consumption,  $MRS_{c_1, c_2}$ , is given by:

$$MRS_{c_1, c_2} = - \left. \frac{dc_2}{dc_1} \right|_P = \frac{\partial P / \partial c_1}{\partial P / \partial c_2} \quad (9)$$

Where a social indifference curve passes through the point where consumption is the same in each period, the curve must be steeper than a downward sloping 45 degree line, which has an absolute slope of 1. This is because time preference implies that the social planner is prepared to give up one unit in the second period in order to get less than one extra unit in

the first period. Hence:

$$MRS_{c_1, c_2} |_{c_1=c_2} = \frac{\partial P / \partial c_1}{\partial P / \partial c_2} > 1 \quad (10)$$

A precise measure of pure time preference can be based on the extent to which the absolute slope of the social indifference curve at  $c_1 = c_2$  exceeds 1, as follows. Suppose the evaluation function  $P$  is additively separable, so that  $P(c_1, c_2) = P_1(c_1) + P_2(c_2)$ . In the case where  $c_1 = c_2 = c$  and consumption is the same in both periods, time preference implies that  $P_1(c) > P_2(c)$ . Writing  $P_1(c) = U(c)$ , it must be possible to write  $P(c, c) = U(c) + \gamma U(c)$ , where  $\gamma < 1$ , and hence when  $c_1 = c_2 = c$ :

$$MRS_{c_1, c_2} |_{c_1=c_2} = \frac{1}{\gamma} \quad (11)$$

To express the fact that  $\gamma < 1$ , write  $\frac{1}{\gamma} = 1 + \rho$ . Clearly  $\rho$  reflects the extent to which the social indifference curve at  $c_1 = c_2 = c$  is steeper than 45 degrees. Hence  $\rho$  measures the rate of pure time preference of the social planner, and  $\gamma = 1 / (1 + \rho)$ .

In general it can be shown that if  $P(C_0) > P(C_1)$ , for two streams  $C_0$  and  $C_1$ , then it is possible to write:

$$\sum_{t=1} \left( \frac{1}{1 + \rho} \right)^{t-1} U(c_{0,t}) \geq \sum_{t=1} \left( \frac{1}{1 + \rho} \right)^{t-1} U(c_{1,t}) \quad (12)$$

where, as above,  $U(c_t)$  represents an evaluation function defined over a single period,  $t$ , in contrast with the multi-period  $P$ . Hence, the ranking according to  $P(C)$  is the same as the ranking according to:

$$W(C) = \sum_{t=1} \left( \frac{1}{1 + \rho} \right)^{t-1} U(c_t) \quad (13)$$

The evaluation function  $W(C)$  has the same form as the welfare function in (1) above. The difference is that in the latter case, pure time preference is simply assumed to be a feature of the social planner, who uses the cardinal weighting function  $U(c)$  in each period: it is necessarily cardinal because the values are added in (13). However, following Koopman's axiomatic approach, time preference is seen to be implied by a set of basic axioms, where evaluation of a time stream is based on an ordinal evaluation function,  $P$ .

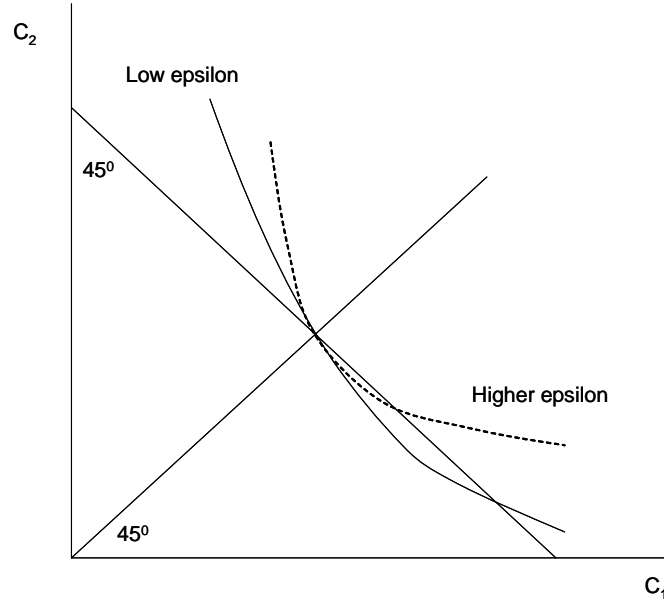


Figure 8: Social Indifference Curves and Time Preference

In general, the absolute slope of a social indifference curve associated with the social welfare function in (13) is:

$$MRS_{c_1, c_2} = - \left. \frac{dc_2}{dc_1} \right|_W = \left( \frac{1}{\gamma} \right) \frac{\partial U / \partial c_1}{\partial U / \partial c_2} = (1 + \rho) \frac{\partial U / \partial c_1}{\partial U / \partial c_2} \quad (14)$$

This may be compared with the special case of an iso-elastic weighting function, where:

$$MRS_{c_1, c_2} = \left( \frac{c_2}{c_1} \right)^\varepsilon (1 + \rho) \quad (15)$$

Examples of two indifference curves are shown in Figure 8. Consumption in each period is equal along the upward sloping 45 degree line from the origin. If, along this ray, each indifference curve has a downward slope of 45 degrees, it means that an increase in period 1's consumption by one unit must be matched by precisely the same fall in period 2's consumption. This would reflect an absence of pure time preference. A positive time preference means that giving up one unit in period 2 requires an increase in period 1 of less than one unit: hence the associated indifference curve is steeper than a downward sloping 45 degree line. In Figure 8, each indifference curve shown



is steeper than the downward sloping 45 degree line. The degree of pure time preference is thus reflected in the difference between the slope of the indifference curve and the 45 degree line.

It can also be seen that the elasticity of marginal valuation, reflecting the concavity of  $U$  (the extent to which the slope, the marginal valuation, falls as consumption increases), is also a measure of the convexity of indifference curves. The solid curve in Figure 8 reflects a lower value of  $\varepsilon$  than the broken curve. The welfare function is also homothetic, whereby the marginal rate of substitution depends on only the ratio of consumption levels (the slopes of indifference curves are the same along any ray from the origin).

## 4 Choice of Optimal Time Stream

Instead of comparing given time-streams, consider a planner, with value judgements represented by the social welfare function in (1), who must decide on the optimal consumption and saving path of the economy. The welfare function is maximised subject to an intertemporal budget constraint which can be written in the form:

$$\sum_{t=1}^T \left( \frac{1}{1+r} \right)^{t-1} c_t = Y \quad (16)$$

where  $Y$  represents a measure of the present value of resources available for consumption over the period, and  $r$  is the rate of interest in a perfect capital market. The Lagrangean for this problem is:

$$L = W + \lambda \left\{ Y - \sum_{t=1}^T \left( \frac{1}{1+r} \right)^{t-1} c_t \right\} \quad (17)$$

Hence first-order conditions, for  $t = 1, \dots, T$ , are:

$$\frac{\partial L}{\partial c_t} = \left( \frac{1}{1+r} \right)^{t-1} \frac{dU}{dc_t} - \lambda \left( \frac{1}{1+r} \right)^{t-1} \quad (18)$$

so that for two periods  $t$  and  $t+1$ :

$$(1+r) \frac{dU/dc_t}{dU/dc_{t+1}} = 1+r \quad (19)$$

Convenient analytical results can be obtained where  $U$  again takes the iso-elastic form  $U(c) = \frac{c^{1-\varepsilon}}{1-\varepsilon}$ , discussed above, so that the absolute value of the elasticity of marginal valuation,  $\left(\frac{c}{dU/dc}\right) \frac{d}{dc} \left(\frac{dU}{dc}\right)$ , is constant and equal to  $\varepsilon$ . Hence (19) becomes:

$$(1 + \rho) \left(\frac{c_{t+1}}{c_t}\right)^\varepsilon = 1 + r \quad (20)$$

Defining  $g_t = \frac{c_{t+1}}{c_t} - 1$ , taking logarithms and using the approximation  $\log(1 + x) = x$ , gives:

$$g_t = \frac{1}{\varepsilon} (r - \rho) \quad (21)$$

This expression is the Euler equation for optimal consumption: it describes the optimal time path of consumption. In this simple problem, if the various rates are constant, consumption either grows or declines at a constant rate, depending on the value of  $r - \rho$ . If the pure time preference rate is equal to the market rate of interest, consumption smoothing is implied, with  $g_t = 0$ .

Rearrangement of (21) gives:

$$r = \rho + \varepsilon g_t \quad (22)$$

This is often referred to the Ramsey equation. It means that at the optimal position, the market rate of interest is equated with  $\rho + \varepsilon g_t$ . It may therefore be said that along the optimal path, the planner equates the marginal return from saving, represented by the market rate of interest,  $r$ , with the marginal cost of saving, represented by  $\rho + \varepsilon g_t$ .

The above analysis of optimal consumption is often used in macroeconomic models of optimal saving; see, for example, Blanchard and Fischer (1989) and Barro and Sala-i-Martin (1995). In such models macroeconomic behaviour is assumed to be captured by the behaviour of a single individual described as a ‘representative agent’, rather than a social planner as discussed here. There is therefore no consideration of aggregation requirements. In some growth models, the representative individual is assumed to be infinitely lived. The introduction of population growth and other complications can produce a different Euler equation from that given in (21), as discussed in section 7 below.

## 5 The Social Time Preference Rate

Previous sections have involved the use of the pure time preference rate,  $\rho$ , of a hypothetical judge in the context of a social welfare function to discount the weighted values  $U(c_t)$  for each period. However, in cost-benefit analyses it is common to compare present values of time streams of money values of consumption, using a ‘consumption discount rate’, rather than the so-called ‘utility discount rate’ (the pure time preference rate),  $\rho$ . Following (22), the consumption discount rate,  $\delta$ , is defined as:

$$\delta = \rho + \varepsilon g_t \quad (23)$$

This rate,  $\delta$ , is also widely referred to as the ‘social time preference rate’. In the context of cost-benefit analyses where money values of an exogenous consumption stream are evaluated, (23) is the fundamental equation that takes a central role when discussing social time preference rates to be used.<sup>10</sup> The social time preference rate can therefore be positive even for  $\rho = 0$ . However, it is possible, if the growth rate of consumption,  $g_t$ , is negative, for  $\delta$  to be negative, if  $\varepsilon$  is relatively large and  $\rho$  is small.

The use of the social time preference rate,  $\delta$ , to evaluation a consumption stream  $c_t$  (rather than the use of  $\rho$  to evaluate a weighted consumption stream  $U(c_t)$ ), involves an evaluation using  $W^*$ , where:

$$W^* = \sum_{t=1}^T c_t \left( \frac{1}{1 + \rho + \varepsilon g_t} \right)^{t-1} \quad (24)$$

It is usually taken for granted that this welfare function gives the same ranking of projects as does the function in (1). However, this is not guaranteed, as can be seen from the following comparisons.<sup>11</sup>

It is convenient to begin with the most favourable case, that is where consumption does in fact grow at a constant proportional rate,  $g_t = g$  for all

---

<sup>10</sup>In this context  $\delta$  does not need to be set equal to the market rate of interest. This contrasts with determination of the optimal growth path, as in the previous section, where  $\delta$  must be equal to the market rate of interest,  $r$ . The latter is determined by, for example, the marginal product of capital – depending on the precise nature of the growth model considered.

<sup>11</sup>For a more detailed treatment, see Creedy (2007)

$t$ . Hence  $c_t = c_1 (1 + g)^{t-1}$  and substitution into (1) gives:

$$W = \sum_{t=1}^T \frac{\{c_1 (1 + g)^{t-1}\}^{1-\varepsilon}}{1 - \varepsilon} \left( \frac{1}{1 + \rho} \right)^{t-1} \quad (25)$$

Rearrangement of this expression gives:<sup>12</sup>

$$W = \frac{c_1^{-\varepsilon}}{1 - \varepsilon} \sum_{t=1}^T c_t \left( \frac{1}{1 + \rho + \varepsilon g} \right)^{t-1} \quad (26)$$

and:

$$W = \frac{c_1^{-\varepsilon}}{1 - \varepsilon} W^* \quad (27)$$

This final result demonstrates that  $W^*$ , obtained by discounting money values of consumption at the social time preference rate, does not necessarily coincide with  $W$ , obtained by discounting  $U(c_t)$  at the pure time preference rate  $\rho$ . For given  $\varepsilon$ ,  $W^*$  automatically gives the same ranking as  $W$  only if  $\varepsilon < 1$  and two consumption streams, with different growth rates, have the same initial value of consumption. Otherwise, inconsistencies can arise. Hence it is advisable to use the basic form of welfare function in (1), with an explicit form for  $U(c_t)$ , rather than discounting the stream  $c_t$  using the rate  $\rho + \varepsilon g$ .

## 6 Time Varying Discount Rates

Previous sections have concentrated on the use of constant discount rates. The expression for the social time preference rate allows for some variation via differences in the growth rate, but it is usual to impose a constant rate. This section considers two cases which give rise to the use of a discount rate which decreases as the time period increases. In subsection 6.1, an alternative approach to specifying the welfare function is examined, while subsection 6.2 briefly discusses uncertainty.

---

<sup>12</sup>This also uses the approximation  $(1 + \rho)(1 + g)^\varepsilon = 1 + \rho + \varepsilon g$ .

## 6.1 An Alternative Welfare Function

Attention has so far been restricted to examining value judgements which can be summarised by the social welfare function in (1), in which weighted consumption values,  $U(c_t)$ , are discounted using a pure time preference rate. When considering an axiomatic derivation of time preference in section 3, the importance of the sensitivity axiom was stressed.<sup>13</sup> This ensures that the infinite future is not allowed completely to dominate the present. With this kind of welfare function, there is a basic tension in that, for very long periods, discounting implies that the present dominates the future. Given that the role of the economist is to examine the implications of adopting alternative value judgements, there is clearly a challenge to specify an alternative formulation of a social welfare function – one which does not imply dominance by either the present or a far distant future period.

Ramsey (1928) realised that without discounting, infinite streams would be non-convergent and therefore could not be ordered using a welfare function like (1). His solution was to measure utility over time as a cumulative sum of the distance from a ‘bliss’ level of utility, but the main problem with this approach is the arbitrariness of the level of ‘bliss’. An alternative value judgement, which has received much attention in the literature concerning income inequality, is a variation of the maxi-min rule discussed by Rawls (1971). In the present context this form of welfare function selects the alternative which maximises the value of the lowest time-stream of consumption. But this criterion also fails to rank all the other streams. Similarly, an objective function that ranks the satisfaction of basic needs above all other outcomes fails to rank other outcomes.

One way of achieving a partial ordering of infinite utility streams without discounting is the overtaking criterion suggested by von Weizacker (1965). According to this welfare function, utility stream A is preferred to B if, after

---

<sup>13</sup>As stressed above, the axiomatic approach does not pretend to prove that time preference must exist, but rather clarifies the value judgements involved. As pointed out by Heal (1998), the independence axiom implies that the trade-off between events today and events in the distant future is independent of what happens in the meantime: not everyone would take this view.

some finite time period,  $T$ , the cumulative utility of stream A is greater than stream B for all time  $t > T$ . However this is only a partial ordering of utility streams because one stream may oscillate above and below another stream indefinitely – it may never permanently overtake. Also, rather than replacing the need for discounting, the overtaking criterion comes close to Koopmans’s axiomatic defence of discounting, because it implies that, for example, utility stream A:  $\{0, 1, 0, 0, \dots\}$  is preferred to stream B:  $\{0, 0, 1, 0, 0, \dots\}$ . Stream B is stream A lagged one period. Hence stream A ‘overtakes’ stream B in period 1 but is identical thereafter. The preference for the overtaking stream therefore reflects a time preference; see also Heal (1998) who made this point.

An alternative form of social welfare function was proposed by Chichilnisky (1997), who was concerned by the fact that choice of a constant discount rate versus a zero rate involves a choice between a ‘dictatorship of the present’ versus a ‘dictatorship of the future’, referred to above. She suggests a solution based on axioms that aim for equity and efficiency in intertemporal resource allocation (the axioms are not discussed in detail here). Her approach is to take an arbitrarily very long time horizon and assign weights to future utility that decline over time and then to assign some extra weight to the utility in the last period. The social welfare function then consists of a weighted average of two terms: the sum of discounted utility where the discount rate declines over time, and the (undiscounted) utility in the final period. Hence:<sup>14</sup>

$$W = \theta \sum_{t=1}^T (1 + \rho_t)^{1-t} U_t(C_t) + (1 - \theta) U_T(C_T) \quad (28)$$

with  $\rho'_t < 0$  and  $0 < \theta < 1$ . Chichilnisky (1997, p. 468) describes the value judgements reflected in this form of social welfare function as ‘sustainable preferences’ because neither the present nor the future is favoured over

---

<sup>14</sup>Chichilnisky’s function differs from (28) in that in her model utility is derived from both consumption,  $c$ , and a flow of services from the stock of natural capital,  $s$ . This implies an optimal combination of  $c$  and  $s$  at any time  $t$ . However, dropping  $s$  from the utility function, as here, does not affect the notion of sustainable preferences, the key ingredients of which are a declining discount rate applied to  $U(\cdot)$  and the second additive term in (28).

the other.<sup>15</sup> This welfare function therefore involves choice of an additional parameter, the weight  $\theta$ .

In justifying the declining discount rate in the first term in (28) Chichilnisky refers to a large body of experimental evidence that the relative weight that people give to two subsequent periods in the future is inversely related to their distance from today.<sup>16</sup> Such preferences have been called hyperbolic preferences (Laibson, 1997). In addition, a hyperbolic discount rate is found to be a necessary condition for an optimal path to exist in maximising (28) in the limit as  $T$  approaches infinity – a constant discount rate would not yield a solution. The second term in (28) justifies the term ‘sustainable preferences’ because it gives explicit recognition to the very long run (at time  $T$ ). Taking a weighted average of the two terms implies a trade-off between the present and the future, yet neither need dominate completely.

These preferences are subject to the standard criticism of time inconsistency that applies to hyperbolic preferences since these are reflected in the first term in (28). But this criticism is weak when the objective is a socially optimal consumption path. From a social choice perspective there is nothing natural or desirable about time consistency, a point made forcefully by Heal (1998). His point is that as new generations arrive and older ones drop out of the choice process, there is no reason why the preferences of generations who have dropped out should be imposed on new generations in the name of time consistency. Chichilnisky’s sustainable preferences cannot of course resolve the debate because they still imply a trade-off between the present and the future which is fundamentally an ethical judgement.

A variation on this type of welfare function reflecting sustainable preferences was introduced by Li and Lofgren (2000). Briefly, society is assumed to be composed of two representative individuals who have utility functions specified over a very long period which spans several generations, on the

---

<sup>15</sup>She cites Solow’s (1992) term ‘intertemporally equitable preferences’ as an alternative description.

<sup>16</sup>To give a simple example, if a person prefers to receive \$90 today than \$100 in one year’s time, but prefers to receive \$100 in 5 years’ time rather than \$90 in 4 years’ time, then their discount rate from year 4 to 5 is lower than that from year zero to 1. Of course, the finding that some people have such preferences does not imply that they *should* be used in social evaluations.

argument that individuals care about their offspring. One individual discounts the future at a constant rate and the other does not discount. The social welfare function is specified as a weighted average of the utilities of the two individuals, with the weights depending on value judgements. It can be shown that the implied time preference rate in this welfare function declines over time.<sup>17</sup> This approach also differs from the welfare function discussed in Sections 3 to 5, where per capita consumption in each period is subject to a weight reflecting the independent judge's value judgements (involving the elasticity of marginal valuation). The Li and Lofgren approach involves an assumed 'representative' form of utility function and the constant pure time preference rate also reflects the properties of the individual who discounts, rather than the judge. The earlier approach took the view that, however individuals in society may actually discount the future in their private consumption decisions, and whatever the degree of concavity of their utility functions, the welfare function embodies only value judgements. The approach here allows the preferences of individuals in the society to carry much more weight.

The advantage of the social welfare specifications discussed here is that they do offer alternative forms in which 'sustainable preferences' arise, in contrast to the extremes arising from (1), with either  $\rho > 0$  or  $\rho = 0$ , in which either the present or the future dominates when very long periods are considered. Consistent with the basic approach recommended in this paper, involving the comparison of alternative value judgements rather than an attempt to impose investigators' own values, the availability of specifications allowing for a wider range of views is to be welcomed.

## 6.2 Uncertainty

The models of Chichilnisky and Li and Lofgren generate a declining consumption discount rate through a declining pure rate of time preference. Alternatively, a declining consumption discount rate could be imposed from the outset as an ad hoc way of taking account of uncertainty regarding the

---

<sup>17</sup>For further details see Li and Lofgren (2000) and the modified welfare function examined in Creedy and Guest (2007).



growth rate; see Weitzman (2007). Suppose the growth rate of consumption is assumed to be constant, but uncertain, giving rise to a distribution of possible values of the consumption discount rate,  $\delta$ . If the possible values are  $\delta_i$ , for  $i = 1, \dots, D$ , and they have associated probabilities of  $p_i$ , Weitzman (2007) showed that the effective consumption discount rate,  $\delta'_t$ , applying to period  $t$  is obtained not as a simple weighted average of the  $\delta_i$ s but of the discount factors. Thus the required relationship is:

$$(1 + \delta'_t)^{-t} = \sum_{i=1}^D (1 + p_i \delta_i)^{-t} \quad (29)$$

so that, by taking logarithms and using the approximation,  $\log(1 + x) = x$ :

$$\delta'_t = -\frac{1}{t} \log \sum_{i=1}^D (p_i \delta_i)^{-t} \quad (30)$$

It is clear from (30) that this declining rate weakens the ‘dictatorship of the present’ that arises with a constant discount rate, as do the models of Chichilnisky and Li and Lofgren.<sup>18</sup>

## 7 The Choice of Unit of Analysis

The previous discussion has assumed that there are no relevant non-income differences between individuals and that population size is constant. Suppose instead that the number of individuals at time  $t$  is  $N_t$  and that individuals of age  $i$  have an equivalent adult size of  $s_i$ , for example because they may have special age-related needs.<sup>19</sup> The equivalent size of the population at  $t$  is  $P_t = \sum_i s_i N_{i,t}$  and the average equivalent size is  $\bar{s}_t = P_t // N_t$ . The question then arises as to the variable, or ‘welfare metric’ to enter the social welfare

---

<sup>18</sup>Weitzman actually gave the continuous time version of the above result. He also discussed the use of certainty equivalent discount rates and the implications of what Arrow and Hurwitz (1972) called ‘pure uncertainty’. On alternative approaches to uncertainty see, for example, Shilizzi (2007) and Woodward and Bishop (1997).

<sup>19</sup>The discussion in this section assumes that population size is exogenous. However, taking a wider perspective means that judgements about population size and composition are themselves involved in social welfare comparisons. This raises important issues that are beyond the scope of the present paper.

function. One approach is to write  $U$ , the weighting function, as a function of the ratio of average consumption to average equivalent size,  $\bar{c}_t/\bar{s}_t = C_t/P_t$ , where  $C_t$  denotes aggregate consumption in period  $t$ . It should be recognised that this is not equal to average consumption per equivalent person, the average value of  $c/s$  in the population at year  $t$ .<sup>20</sup>

Given a distinction between individuals and equivalent persons, a further decision must be made about the unit of analysis in a welfare function. This decision again involves value judgements. The question of choice of units has been considered in the literature on inequality measurement, but has received little attention in multi-period contexts; for an exception, see Creedy and Guest (2008).<sup>21</sup> Statements about comparisons between households, in the context of inequality, can easily be converted to statements about comparisons between time periods.

One approach to defining a unit of analysis is to use the ‘adult equivalent person’. In the multiperiod context, there are  $P_t$  adult equivalent persons at time  $t$ , and so the social welfare function becomes:

$$W = \sum_{t=1} P_t \left( \frac{C_t}{P_t} \right)^{1-\varepsilon} \frac{(1+\rho)^{t-1}}{1-\varepsilon} \quad (31)$$

The resulting Euler equation for optimal growth at  $t$  is:

$$g_t = \frac{1}{\varepsilon} (r - \rho) + (p_t - n_t) \quad (32)$$

where  $g_t$ ,  $p_t$  and  $n_t$  are respectively the proportional rates of change of  $C_t$ ,  $P_t$  and  $N_t$ . In this way, the ‘income’ concept and the unit of analysis are treated consistently, ensuring that each individual’s contribution depends on the demographic structure of the time period to which they belong. An alternative approach is to treat the individual as the basic unit of analysis.

---

<sup>20</sup>The two terms are equal either if  $c_{i,t}/s_{i,t}$  is constant for all  $i$ , or if  $s_{i,t}$  and  $c_{i,t}$  are uncorrelated.

<sup>21</sup>Major contributions in the context of inequality include Shorrocks (2004), Decoster and Ooghe (2002), Glewwe (1991) and Ebert (1997). The use of different units can lead to opposite conclusions about the effects on inequality of a tax policy change. Examples of such conflicts using tax microsimulation models are given by Decoster and Ooghe (2002) and Creedy and Scutella (2004).

As there are  $N_t$  individuals at time  $t$ , the social welfare function can be written as:

$$W = \sum_{t=1} N_t \left( \frac{C_t}{P_t} \right)^{1-\varepsilon} \frac{(1+\rho)^{t-1}}{1-\varepsilon} \quad (33)$$

For the optimal consumption path problem, the Euler equation is found to be:

$$g_t = \frac{1}{\varepsilon} \{r - \rho + (\varepsilon - 1)(p_t - n_t)\} \quad (34)$$

so that although the difference between the social welfare functions (31) and (33) concerns only the choice of weights in each period, that is a choice between  $P_t$  or  $N_t$ , the resulting optimal consumption paths can differ substantially. This is because the choice between individuals and adult equivalents as the basic unit of analysis can in principle lead to different conclusions about the effects of transferring consumption between time periods, which has implications for the path of optimal consumption.

## 8 Measurement Attempts

It has been stressed that the elasticity of marginal valuation,  $\varepsilon$ , is not an objective measure relating to individuals in society, but reflects the subjective value judgements of a fictional judge who is evaluating alternative policies or outcomes. However, a superficially similar-looking concept of the ‘elasticity of marginal utility’ plays a central role in some consumer demand systems, particularly where directly additive utility functions are involved; this was first clarified by Frisch (1959). Define the elasticity of the marginal utility of total expenditure with respect to total expenditure as  $\xi$ . If  $\delta_{\ell j}$  denotes the Kroneker delta, such that  $\delta_{\ell j} = 0$  when  $\ell \neq j$ , and  $\delta_{\ell j} = 1$  when  $\ell = j$ , and  $e_\ell$  is the total expenditure elasticity for good  $\ell$ , and  $w_j$  is the budget share of good  $j$ , Frisch showed that the price elasticities,  $e_{\ell j}$ , can be written as:

$$e_{\ell j} = -e_\ell w_j \left( 1 + \frac{e_j}{\xi} \right) + \frac{e_\ell \delta_{\ell j}}{\xi} \quad (35)$$

This can be used to obtain  $\xi$ , given independent values of the elasticities. In the special case of the Linear Expenditure System, the elasticity of marginal

utility has a convenient interpretation: it is the ratio of total expenditure to supernumerary expenditure, that is, expenditure above a ‘committed’ amount. The use of the LES in empirical demand studies therefore necessarily involves an (absolute) elasticity value which is well above unity, and studies typically obtain a value of around 2.

In view of the entirely different contexts of welfare comparisons involving social evaluation functions and empirical studies of household consumption behaviour, there is no relationship whatsoever between  $\varepsilon$  and  $\xi$ . In other words, there is no reason why a value of  $\varepsilon$ , to be imposed in making comparisons, could be ‘estimated’ using information from studies of household budgets. Nevertheless, this suggestion is sometimes made; for example, elasticities obtained on the basis of the Linear Expenditure System are discussed by Evans (2005, pp. 204-206).

In considering alternative values of  $\varepsilon$ , it is useful to ensure that they are within a range that is considered appropriate by potential users of the results: for example there is little point in reporting values of, say,  $\varepsilon > 10$  if the vast majority of readers would regard them as extreme. Hence questionnaire studies have been designed to elicit information about individuals’ value judgements.<sup>22</sup> It was in this spirit that the questionnaire study of Amiel *et al.* (1999) was carried out. Nevertheless, there was no suggestion that questionnaires can produce any single value that should be used in policy evaluations.<sup>23</sup> A substantial number of respondents did not adhere to the constant relative inequality aversion form. In addition, Amiel and Cowell (1994) have found that a large number of questionnaire respondents do not actually share the value judgements that are explicit in the most common forms of social welfare function used in evaluation work, such as the one discussed above. This presents a challenge to produce alternative flexible specifications.

---

<sup>22</sup>Early questionnaire studies were carried out by Glesjer *et al.* (1977), and Gevers *et al.* (1979), although no attempt was made to estimate precise specifications of distributional preferences.

<sup>23</sup>There is sometimes confusion about this, as in Evans (2005).

## 8.1 Taxation and Equal Absolute Sacrifice

A different approach to the ‘estimation’ of  $\varepsilon$  involves attempts to estimate the implicit value judgements revealed by tax and transfer policies. Some authors, including Stern (1977), Cowell and Gardiner (1999) and Evans (2005), have suggested that such estimates provide a guide to  $\varepsilon$  values which should be applied in policy evaluations.<sup>24</sup>

There are two steps to such an approach. The first step attempts to infer, from tax policy decisions, value judgements which are not otherwise made explicit. Such an attempt can be defended – provided of course that the model used to make inferences is plausible. For example, such estimates may be useful in checking whether there is in fact any correspondence between policies and basic value judgements of policy makers. Given the complexities involved in tax policy design, it may be useful to know if a particular structure is associated with implicit judgements that may actually be very different from those held (though seldom made explicit).<sup>25</sup>

The second step is the illegitimate one of suggesting that estimates of implicit value judgements ‘should’ be used in making social evaluations. This criticism applies even in the most unlikely case where implicit views can be identified precisely.

This subsection considers the first ‘positive’ step taken by the authors mentioned above, and suggests that the model used is inadequate. The approach is based on the assumption that income tax policy-makers aim to achieve equal absolute sacrifice. It assumes that incomes are exogenously given, rather than arising from endogenous labour supply behaviour (subject to endowments and education which give rise to individual productivities). Suppose  $x$  represents income and the tax function is  $T(x)$ . Equal absolute sacrifice requires, for all  $x$ , that the absolute difference between pre-tax and post-tax utility is the same for all individuals. Hence:

$$U(x) - U(x - T(x)) = k \tag{36}$$

---

<sup>24</sup>Other attempts to infer value judgements include Brent (1984), Christiansen and Jansen (1978) Mera (1969) and Moreh (1981).

<sup>25</sup>This is the view taken by van de Ven and Creedy (2005) when examining adult equivalence scales implicit in tax and transfer systems.

where  $U(\cdot)$  represents a utility function which is considered to be the same for all individuals. The parameter  $k$  depends on the amount of revenue per person.<sup>26</sup> The combination of equal absolute sacrifice with the iso-elastic function,  $U(x) = x^{1-\varepsilon_\tau} / (1 - \varepsilon_\tau)$  for  $\varepsilon_\tau \neq 1$ , gives, from (36) above:<sup>27</sup>

$$\frac{x^{1-\varepsilon_\tau}}{1 - \varepsilon_\tau} - \frac{(x - T(x))^{1-\varepsilon_\tau}}{1 - \varepsilon_\tau} = k \quad (37)$$

Differentiation and simplification gives, as in Evans (2005, p.207), the result that:

$$\log(1 - T'(x)) = \varepsilon_\tau \log\left(1 - \frac{T(x)}{x}\right) \quad (38)$$

where  $T'(x)$  and  $T(x)/x$  are marginal and average tax rates. This expression has been used to carry out ordinary least squares regressions using income tax schedules, so that  $\varepsilon_\tau$  and its standard error are obtained as a regression coefficient. There is some difference of opinion over whether to include a constant in the regression: compare Cowell and Gardiner (1999) and Evans (2005), who also use different income measures. Alternatively, (38) can be rearranged to get  $\varepsilon_\tau = \log(1 - MTR) / \log(1 - ATR)$ , and ‘estimates’ of  $\varepsilon_\tau$  are obtained and compared using simply the marginal and average tax rates at different income levels.

This approach automatically produces a value of  $\varepsilon_\tau$  in excess of unity for a progressive tax system, for which the marginal tax rate exceeds the average tax rate. This feature was first discussed by Edgeworth (1897) and formally shown by Samuelson (1947). The values of  $\varepsilon_\tau$  obtained in this way are thus severely constrained by the specification of the objective of equal absolute sacrifice. Furthermore, those using the approach to ‘estimate’  $\varepsilon_\tau$  ignore the objections raised by Edgeworth and others concerning the various interpretations of sacrifice theories.<sup>28</sup>

<sup>26</sup>This differs from an alternative view that would replace  $U(x)$  with  $W(x)$ . Thus, as with inequality measurement, a judgement is made regarding the welfare metric, and then a view is taken about variations in  $x$ . This judgement is quite separate from the way individuals may themselves view such variations.

<sup>27</sup>Young (1987) actually showed that the iso-elastic form is required if an indexation requirement is imposed on the tax structure in addition to equal sacrifice. But of course fiscal drag is a common, indeed almost universal, feature of income tax structures.

<sup>28</sup>This does not apply to those, such as Richter (1983) and Young (1987) who were interested only in deriving the implications of various axioms.

For  $\varepsilon_\tau > 1$ , equation (37) can be rearranged as:<sup>29</sup>

$$T(x) = x - \{x^{1-\varepsilon_\tau} - k(1 - \varepsilon_\tau)\}^{1/(1-\varepsilon_\tau)} \quad (39)$$

which gives smooth and increasing marginal and average rate schedules. Of course, in practice tax functions are multi-step functions with ranges where the marginal rate is constant. In some structures there is a ‘standard rate’ which applies over a wide range of taxable income, so the above function obviously has difficulty capturing this range.<sup>30</sup> The imposition of  $\varepsilon > 1$  is highly restrictive. Furthermore, the model applies only to positive taxes. It can thus relate at best to a small component of a much broader set of taxes and transfers.

Hence some scepticism must be attached to interpretations of estimates obtained using this model as implicit value judgements. It seems most likely that the approach has been chosen largely – or indeed only – for its simplicity. On the other hand, the optimal tax framework has demonstrated the considerable complexity involved in the link between value judgements and the tax structure and, importantly, progression can arise with values of  $\varepsilon < 1$ . But of course even if the estimation of implicit preferences were considered plausible, they cannot qualify as value judgements which should be imposed. There is no alternative to accepting that value judgements are required and the best attitude of professional economists is to report a range of results based on alternative value judgements. In reporting results, readers need to appreciate precisely what is implied about value judgements by different values of  $\varepsilon$ , since it is not immediately obvious whether, for example, a value of  $\varepsilon = 0.5$  indicates a high or low aversion to inequality. This is considered

---

<sup>29</sup>In stating this result, Young (1987, p. 212) rewrote  $-k(1 - \varepsilon)$  as  $\lambda^{1-\varepsilon}$ , so that the tax function compares with a constant elasticity of substitution form.

<sup>30</sup>If equal absolute sacrifice is combined with a welfare function displaying constant absolute inequality aversion,  $\alpha$ , such that  $W = 1 - \exp(-\alpha x)$ , a tax function of the form  $T(x) = x + \frac{1}{\alpha} \log \{k + e^{-\alpha x}\}$  arise. This can be made to display rate schedules similar to those illustrated above. Dalton (1954, pp.68-70) discussed several examples using alternative utility functions and sacrifice principles, and showed that if equal absolute sacrifice produces progression, equal proportional sacrifice produces a more progressive tax structure. In an early study, Preinreich (1948) considered the form of the utility schedule consistent with the US tax legislation, without imposing a specific functional form over the whole income range. He assumed equal proportional sacrifice.

in the following section.

## 9 Interpreting Orders of Magnitude

In using  $\varepsilon$  values to compute values of social welfare functions, or carry out cost-benefit evaluations, there is no alternative but to consider a range alternative values, implying different degrees of aversion to inequality. In some cases, ‘dominance’ results may be obtained. In other words one policy may be judged to give rise to a higher value of social welfare than another policy for all values of  $\varepsilon$ . In other situations, readers can make up their own minds given the reported computations. It is therefore important to appreciate the precise nature of the comparisons being made. When the link between a social welfare function and a measure of inequality was introduced by Atkinson (1970), he recognised the difficulty of forming views about the orders of magnitude of  $\varepsilon$  using the welfare function  $W = \sum_{h=1}^H \frac{y_h^{1-\varepsilon}}{1-\varepsilon}$ . In order to help interpretation, he used the idea of a ‘leaky bucket’ experiment, which considers the extent to which a judge is prepared to tolerate some loss in making a transfer from one person to another.<sup>31</sup>

Consider two individuals, so that from the welfare function, setting the total differential equal to zero gives:

$$-\left. \frac{dy_1}{dy_2} \right|_W = \left( \frac{y_1}{y_2} \right)^\varepsilon \quad (40)$$

The welfare function is thus homothetic, as the slopes of social indifference curves are the same along any ray drawn through the origin. Consider two individuals and, using discrete changes, suppose a dollar is taken from the richest, such that  $\Delta y_2 = -1$ . The amount to be given to the other individual to keep social welfare unchanged is thus:

$$\Delta y_1 = \left( \frac{y_1}{y_2} \right)^\varepsilon \quad (41)$$

For example, if  $y_2 = 2y_1$  and  $\varepsilon = 1.5$ , it is necessary to give person 1 only

---

<sup>31</sup>Okun (1975) examined a slightly different kind of leaky bucket experiment involving transfers between groups of individuals.



35 cents – a leak of 65 cents from the original dollar taken from person 2 is tolerated. If  $\varepsilon = 1$ , a leak of 50 cents is tolerated.

This type of experiment, and thus the sensitivity of the tolerance for a leaking bucket, is well-known in the literature on inequality measurement. But in other contexts in which the same kind of iso-elastic function is used, relatively large values of  $\varepsilon$  are often adopted without, it seems, consideration of such implications.<sup>32</sup> For example, in the intertemporal literature, a value of  $\varepsilon = 2$  is often used. Suppose that total income (or consumption) in the first period is 100 and this grows at a rate of 0.02 per period. In period 10 it is thus 119.5, and a judge with  $\varepsilon = 2$  would be prepared to take a dollar from period 10, and give only \$0.70 to period 1. By period 20 total income would be 145.7, and the same judge would reduce period 20's income by \$1 while adding only \$0.47 to the first period. The social time preference rate is thereby increased significantly above the pure time preference rate. The leaky bucket experiment therefore provides a useful illustration of the implications, in terms of value judgements, of adopting particular values of  $\varepsilon$  in any policy evaluation.

## 10 Conclusions

This paper has provided an analytical review of problems arising in the evaluation of alternative time streams of consumption using the concept of time preference. The potential sensitivity of comparisons, especially to the choice of time preference rate and elasticity of marginal valuation, was stressed. The nature of time preference, based on an axiomatic approach, was examined. The analysis of individual optimisation over time then led to the concept of the social time preference rate, and a difficulty with using this rate was highlighted and complications introduced by non-income differences between individuals were examined. Attempts to measure the elasticity of marginal valuation were critically discussed.

A basic theme of this paper is that ultimately, evaluations cannot avoid value judgments, so the role of the economist is to examine the implications

---

<sup>32</sup>However, it is discussed by Pearce and Ulph (1998, pp.280-281).

of adopting alternative value judgements. As argued by Varian (2006), ‘Exploring the implications of alternative assumptions is likely to lead to better policy than making a single blanket recommendation. At least at this stage of our understanding, exploration beats exhortation’.<sup>33</sup> In view of the fact that, in the formulations discussed here, there are few parameters representing value judgements, sensitivity analyses do not present any problems computationally. There is no excuse for computing and presenting only one set of values. However, care is needed regarding presentation of results. First, it is useful to investigate whether ‘dominance’ results can be obtained: that is, is one policy option judged superior to others for all, or at least a very wide range, of parameter value combinations? When clearly specified alternative policies are being considered, it is more useful to report critical values of parameters, and their combinations, for which a particular policy dominates, rather than using an arbitrary range of values. For example if there are just two alternative consumption streams, a diagram showing combinations of  $\varepsilon$  and  $\rho$  (where the relevant social welfare function is involved) which ‘divide’ the two policies can be constructed. Above all, the nature of the comparisons must be explained as clearly as possible to policy makers. Here the type of leaky bucket experiment discussed above can be useful. Above all, economists have a duty to make it clear that there is no observable ‘social time preference rate’ or ‘elasticity of marginal valuation’ which ‘should’ be imposed in any policy analysis.

---

<sup>33</sup>Much earlier, Robbins (1935, p.149) argued that ‘this is not to say that economist may not assume as postulates different judgments of value, and then on the assumption that these are valid enquire what judgment is to be passed upon particular proposals for action.’

## References

- [1] Amiel, Y. and Cowell, F.A. (1994) Income inequality and social welfare. In *Taxation, Poverty and Income Distribution* (ed. by J. Creedy), pp. 193-219. Edward Elgar, Aldershot, UK.
- [2] Amiel, Y., Creedy, J. and Hurn, S. (1999), 'Measuring attitudes towards inequality', *Scandinavian Journal of Economics*, 101, 83-96.
- [3] Arrow, K. and Hurwicz, L. (1972) An optimality criterion for decision-making under ignorance. In *Uncertainty and Expectations: Essays in Honour of G.L.S. Shackle* (Edited by C.F. Carter and J.L. Ford). Basil Blackwell, Oxford, UK.
- [4] Atkinson, A.B. (1970), 'On the measurement of inequality', *Journal of Economic Theory*, 2, 244-263.
- [5] Barro and Sala-i-Martin (1995) *Economic Growth*. McGraw-Hill, New York, US.
- [6] Blanchard and Fischer (1989) *Lectures in Macroeconomics*. MIT Press, Cambridge, Mass, UK.
- [7] Brent, R.J. (1984), 'On the use of distributional weights in cost-benefit analysis: a survey of schools', *Public Finance Quarterly*, 12, 213-230.
- [8] Caplin, A. and Leahy, J.V. (2000), 'The social discount rate', *NBER Working Paper* no. W7983.
- [9] Carter, R.M. *et al.* (2006), 'The Stern review: a dual critique', *World Economics*, 7, 165-232.
- [10] Chichilnisky, G. (1997), 'What is sustainable development?', *Land Economics*, 73, 467-91.
- [11] Christiansen, V. and Jansen, E.S. (1978), 'Implicit social preferences in the Norwegian system of social preferences', *Journal of Public Economics*, 10, 217-245.

- [12] Cowell, F.A. and Gardiner, K. (1999) Welfare weights. STICERD, London School of Economics, UK.
- [13] Creedy, J. (2008), ‘Evaluating policy: welfare weights and value judgements’. *Australian Journal of Labour Economics* (forthcoming).
- [14] Creedy, J. (2007) Sustainable preferences and damage abatement: implications for living standards. *University of Melbourne Department of Economics Working Paper*.
- [15] Creedy, J. and Guest, R. (2008), ‘Population ageing and consumption: representative agent versus social planner’. *Economic Modelling* (forthcoming)
- [16] Creedy, J. and Scutella, R. (2004), ‘The role of the unit of analysis in policy reform evaluations of inequality and social welfare’, *Australian Journal of Labour Economics*, 7, 89-108.
- [17] Dalton, H. (1954) *Public Finance* (4th Edition). Routledge and Kegan Paul, London, UK.
- [18] Dasgupta, P. (2006) Comments on the Stern review of climate change. University of Cambridge, UK.
- [19] Decoster and Ooghe (2002) Weighting with individuals, equivalent individuals or not weighting at all: does it matter empirically?, In *Inequality, Welfare and Poverty: Theory and Measurement* (ed. by Y. Amiel and J.A. Bishop), pp. 173-190. Elsevier, New York, US.
- [20] Ebert, U. (1997), ‘Social welfare when needs differ: an axiomatic approach’, *Economica*, 64, 233-244.
- [21] Edgeworth, F.Y. (1897) The pure theory of taxation, *Economic Journal*, Reprinted in *Papers Relating to Political Economy*, 2, 63-125. Macmillan, London, UK.
- [22] Evans, D.J. (2005), ‘The elasticity of marginal utility of consumption: estimates for 20 OECD countries’, *Fiscal Studies*, 26, 197-224.

- [23] Frisch, R. (1959), ‘A complete scheme for computing all direct and cross demand elasticities in a model with many sectors’. *Econometrica*, 27, 177-196.
- [24] Gevers, L., Glesjer, H. and Rouyer, J. (1979), ‘Professed inequality aversion and its error component’, *Scandinavian Journal of Economics*, 81, 238-243.
- [25] Glesjer, J., Gevers, L., Lambot, Ph. and Morales, J.A. (1977), ‘Professed inequality aversion among students’, *European Economic Review*, 10, 173-188.
- [26] Glewwe (1991), ‘Household equivalence scales and the measurement of inequality: transfers from the poor to the rich could decrease inequality’, *Journal of Public Economics*, 44, 214-216.
- [27] Heal, G. (1998) *Valuing the Future: Economic Theory and Sustainability*. Columbia University Press, New York, US.
- [28] Koopmans, T.C. (1960), ‘Stationary Ordinal Utility and Impatience’, *Econometrica*, 28, 287-309.
- [29] Koopmans, T.C., Diamond, P.A. and Williamson, R.E. (1964), ‘Stationary Utility and Time Perspective’. *Econometrica*, 32, 82-100.
- [30] Laibson, D. (1996), ‘Hyperbolic discount functions, undersaving, and savings policy’, *NBER Working Paper*, no. 5635.
- [31] Li, C-Z. and Lofgren, K-G. (2000), ‘Renewable resources and economic sustainability: a dynamic analysis with heterogenous time preferences’, *Journal of Environmental Economics and Management*, 40, 236-250
- [32] Marini, G. and Scaramozzino, P. (2000), ‘Social time preference’, *Journal of Population Economics*, 13, 639-645.
- [33] Mera, K. (1969), ‘Experimental determination of relative marginal utilities’, *Quarterly Journal of Economics*, 83, 464-477.

- [34] Moreh, J. (1981), 'Income inequality and the social welfare function', *Journal of Economic Studies*, 8, 25-37.
- [35] Nordhaus, W.D. (2006), 'The Stern review on the economics of climate change', *National Bureau of Economic Research Working Paper*, no. 12741.
- [36] Okun, A.M. (1975) *Equality and Efficiency: The Big Trade-off*. The Brookings Institution, Washington, US.
- [37] Padilla, E. (2002), 'Intergenerational equity and sustainability', *Ecological Economics*, 41, 69-83.
- [38] Pearce, D. and Ulph, D. (1998) A social discount rate for the United Kingdom. In *Economics and Environment: Essays on Ecological Economics and Sustainable Development* (ed. by D. Pearce). Edward Elgar, Cheltenham, UK.
- [39] Pigou, A. C. (1932) *The Economics of Welfare*, 4th edition. Macmillan, London, UK.
- [40] Preinreich, A.D. (1948), 'Progressive taxation and sacrifice', *American Economic Review*, 38, 103-117.
- [41] Ramsey, F. P. (1928), 'A mathematical theory of saving', *Economic Journal*, 38, 543-59.
- [42] Rawls, J. (1971) *A Theory of Justice*. Harvard University Press, Harvard, US.
- [43] Richter, W.F (1982), 'From ability to pay to concepts of equal sacrifice', *Journal of Public Economics*, 20, 211-229.
- [44] Robbins, L. (1935) *An Essay on the Nature and Significance of Economic Science* (2nd Edition). Macmillan, London, UK.
- [45] Samuelson, P.A. (1947) *Foundations of Economic Analysis*. Harvard University Press, Cambridge MA, US.

- [46] Shilizzi, S. (2007) Discounting the distant future using short time horizons: investments with irreversible benefits. In *Economics and the Future: Time Discounting in Private and Public Decision Making* (Edited by D. Pannell and S. Shilizzi). Edward Elgar, Cheltenham, UK.
- [47] Shorrocks, A. F. (2004), 'Inequality and welfare evaluation of heterogeneous income distributions', *Journal of Economic Inequality*, 2, 193-218.
- [48] Solow, R. (1992) An almost practical step toward sustainability. Invited lecture on the occasion of the fortieth anniversary of Resources for the Future, 8 October 1992.
- [49] Stern, N. (1977) Welfare weights and the elasticity of the marginal valuation of income. In *Studies in Modern Economic Analysis* (ed. by M. Artis and R. Nobay). Basil Blackwell, Oxford, UK.
- [50] Stern, N. (2006) The Economics of Climate Change. Available at [http://www.hm-treasury.gov.uk/independent\\_reviews/stern\\_review\\_economics\\_climate\\_change/sternreview\\_index.cfm](http://www.hm-treasury.gov.uk/independent_reviews/stern_review_economics_climate_change/sternreview_index.cfm)
- [51] van de Ven, J. and Creedy, J. (2005), 'Taxation, reranking and equivalence scales', *Bulletin of Economic Research*, 57, 13-36.
- [52] Varian, H.L. (2006), 'Recalculating the costs of global climate change', *New York Times* (14 December).
- [53] von Weizacker, C. (1965), 'Existence of Optimal Programs of Accumulation for an Infinite Time Horizon', *Review of Economic Studies*, 32, pp. 85-104.
- [54] Woodward, R.T. and Bishop, R.C. (1997), 'How to decide when experts disagree: uncertainty-based choice rules in environmental policy'. *Land Economics*, 73, 492-507.
- [55] Young, H.P. (1987), 'Progressive taxation and the equal sacrifice principle', *Journal of Public Economics*, 32, 203-214.