

## **Abstract**

The chapter introduces the consumption or social discount rate. In application, the rate is used in cost benefit analysis of public projects and legislation. In theoretical reasoning, this discount rate is a tool to analyze intertemporal trade-offs and the intergenerational weights implicit to general economic frameworks. We introduce the rate in a model assuming a single agent consuming an aggregate commodity with certainty. Then we relax these assumptions. First, we discuss how discounting changes under risk and uncertainty. Second, we consider the case that environmental and produced goods are of limited substitutability and grow at different rates. Third, we discuss a setting of overlapping, potentially altruistic agents. These examples can all lead to non-constant term structures. We discuss how to deal with the particular case where non-constant discount rates trigger time inconsistent plans.

# 1 Introduction

Consumption today is not directly comparable with consumption at a different point in time. The discount factor for consumption enables us to compare consumption at different points in time. Discounting is an especially important element of environmental problems that involve trade-offs in consumption across widely different times. Climate policy is the leading example of this kind of trade-off, because decisions taken in the near future may have major effects on welfare in the distant future.

## 1.1 The social discount rate

Discount rates are defined as the rate of decrease (the negative of the rate of change) of the discount factor. It is important at the outset to distinguish between discount rates and factors for utility and for consumption. We define  $\beta_t$  as the number of units of utility (utils) that we would give up today in order to obtain one more util at time  $t$ . It is the discount factor for utility. By definition,  $\beta_0 = 1$ . We define the discount rate for utility at time  $t$  as<sup>1</sup>

$$\rho_t = -\frac{\dot{\beta}_t}{\beta_t},$$

where the dot denotes the (total) time derivative. The utility discount rate  $\rho_t$  is also known as the rate of pure time preference (RPTP). The RPTP is a measure of impatience, with larger values implying greater impatience. If  $\rho_t = \rho$  is a constant, utility discounting is exponential:  $\beta_t = e^{-\rho t}$ .

We begin by defining the discount factor and the corresponding discount rate for consumption in the simplest case: there is a single consumption good,  $c$ ; there is no uncertainty; and welfare,  $W$ , equals the present value of the infinite stream of utility,  $u(c)$ . In this case,  $W = \int_0^\infty \beta_t u(c_t) dt$ . The consumption discount factor for time  $t$  equals the number of units of consumption we would surrender during a small interval,  $\varepsilon$ , beginning today in order to obtain one more unit of consumption during a small interval,  $\varepsilon$ , beginning at time  $t$ . If, prior to the transfer, consumption today is  $c_0$  and consumption at time  $t$  is  $c_t$ , the welfare loss due to giving up  $\Gamma$  units of

---

<sup>1</sup>Here we define the instantaneous discount rate. Another frequent definition of discount rates is as an average rate defined by  $\frac{1}{t} \ln \beta_t$ .

consumption today is approximately  $u'(c_0)\Gamma\varepsilon$  and the welfare gain of one unit of consumption at time  $t$  is  $\beta_t u'(c_t)\varepsilon$ . We are willing to make this sacrifice if these two quantities are equal, i.e. if

$$\Gamma_t = \beta_t \frac{u'(c_t)}{u'(c_0)}. \quad (1)$$

The rate of decrease of  $\Gamma$  is the discount rate for consumption. This rate is more conveniently expressed in terms of the growth rate of consumption  $g$  and the consumption elasticity of marginal utility  $\eta$ , which is equal to the inverse of the elasticity of intertemporal substitution. These are defined as

$$g_t = \frac{\dot{c}_t}{c_t} \quad \text{and} \quad \eta(c_t) = -\frac{u''(c_t)}{u'(c_t)}c_t.$$

Then, equation (1) gives rise to the consumption discount rate

$$r_t = -\frac{\dot{\Gamma}_t}{\Gamma_t} = \rho_t + \eta(c_t)g_t. \quad (2)$$

Equation (2) is usually referred to as the Ramsey equation. More precisely, the actual Ramsey equation is an equilibrium condition in the Ramsey-Cass-Koopmans growth model stating that the right hand side of equation (2) has to equal the interest rate (or capital productivity) in the economy. In contrast, the derivation of the consumption discount rate  $r_t$  in equation (2) is independent of the market equilibrium. In the context of public project evaluation, the consumption discount rate  $r_t$  is referred to as the social discount rate (SDR).

A larger SDR means that we are willing to sacrifice fewer units of consumption today in order to obtain an additional unit of consumption  $t$  periods in the future. In the context of climate policy, a larger SDR means that we are willing to spend less today, e.g. through increased abatement or investment in low-carbon technology, to prevent future climate damage. A larger value of the RPTP,  $\rho$ , means that we place less value on future utility. A higher growth rate means that future consumption is higher; under the assumption of decreasing marginal utility of consumption the higher  $g$  lowers future marginal utility. A larger elasticity of marginal utility implies a faster decrease of marginal utility with consumption growth. Therefore, under positive growth, larger values of  $\rho$ ,  $g$ , or  $\eta$  all imply a higher SDR, and less concern for the future.

Some applications assume: (i) isoelastic utility  $u(c) = \frac{c^{1-\eta}}{1-\eta}$ , so that  $\eta$  is constant; (ii) a constant growth rate, so that  $g$  is constant; and (iii) exponential discounting of utility, so  $\rho$  is constant. In this case, the SDR is also constant. More generally, one or more of the three components of  $r_t$  might depend on time. While  $g_t$  or  $\eta(c_t)$  will quite commonly depend on time because of the dynamics in the economy, a time dependence of pure time preference would be exogenously imposed as a direct preference specification. As we discuss in section 3.2 such a time dependence of pure time preference can lead to time inconsistent choices.

## 1.2 The SDR and environmental policy

The social discount rate is used to evaluate legislation and public projects. In application, the employed values vary widely over countries and agencies. While the majority adopts a constant rate, the U.K. and France adopt time dependent discounting schemes. The social discount rate is important in evaluating environmental policy when the timing of costs and benefits differ, as with climate change policy where current decisions have long-lasting effects. We use the latter as an example to illustrate the policy relevance of the social discount rate. The Stern (2007) Review of Climate Policy uses a negligible RPTP of  $\rho = 0.1\%$ , a growth rate of  $g = 1.3\%$ , and the value  $\eta = 1$ , implying  $r = 1.4\%$ . In contrast, Nordhaus (2008) employs a RPTP of  $\rho = 1.5\%$  and the value  $\eta = 2$  in a model with an approximate consumption growth rate of  $g = 1.5\%$ , implying  $r = 5.5\%$ . The ratio of what we are willing to spend today, to avoid a dollar of damage 100 years from now, under these two sets of parameters is

$$\frac{\Gamma_{100}^{Stern}}{\Gamma_{100}^{Nord}} = \frac{e^{-.014 \cdot 100}}{e^{-.055 \cdot 100}} \approx 60.$$

For this example, the higher SDR decreases our willingness to spend today to avoid damages a century from now by a factor of 60. Nordhaus (2007) shows that this difference in social discounting can explain almost entirely the differences in policy recommendation between his integrated assessment of climate change based on DICE-2007 and that of the Stern Review: running DICE with a 1.4% SDR instead of 5.5% increases the near term social cost of carbon by a factor of 10 and almost quadruples the near term optimal abatement rate with respect to business as usual.

### 1.3 The positive and the normative perspective

The different choices of the social discount rate in Nordhaus's (2008) and Stern's (2007) assessment of climate change stem from different perspectives on the application of social discounting in policy evaluation. Nordhaus takes a positive approach to social discounting, while Stern takes a normative approach. The positive approach relies on measurements of the constituents of the social discount rate, while the normative approach chooses these parameters on ethical grounds. The measurement of the social discount rate is complicated by the fact that markets exhibit externalities, are incomplete, and, therefore, do not necessarily reflect the agents' intertemporal preferences correctly.

In principle, there are two different approaches to determine the social discount rate as reflected on the market. First, we can measure  $\rho$  and  $\eta$  based on a sufficiently large set of observations. We then combine these estimates with an exogenous consumption growth scenario, or use them in an evaluation model where growth is determined endogenously, as in the integrated assessment of climate change. Second, we can measure the interest rate in the economy. Then, the original derivation of the Ramsey (1928) equation (2) states that in equilibrium this interest rate is equivalent to the consumption discount rate. This second method is particularly prone to picking up market imperfections like transaction costs or distortions in the intertemporal consumption-investment trade-off. These market imperfections also result in a wide spectrum of different interest rates observable on the market. Usually, interest rate based approaches to measuring the social discount rate rely on the interest paid on government bonds. These provide an opportunity cost measure for a dollar spent on public projects. The maturity of government bonds limits how far into the future we can measure this opportunity cost; in the U.S. it is currently given by the 30-year treasury bond.

The Stern (2007) Review argues that intergenerational trade-offs encompassing many generations cannot be judged merely on the basis of market observations. Society has to employ ethical reasoning in order to represent those generations that are not currently alive and, hence, not reflected on the market. The sequence of economists who argued that ethical reasoning imposes a zero RPTP is long and includes Ramsey (1928), Pigou (1932), Harrod (1948), Koopmans (1963), Solow (1974), Broome (1992). While the Stern Review's choice of a close to zero RPTP is based on intergenerational

equity concern, it simultaneously adopts a comparatively low value for the propensity to smooth consumption over time  $\eta$ , implying a rather low preferences for intergenerational consumption smoothing. Traeger (2012a) presents a different normative argument for a zero RPTP based entirely on rationality constraints for decision making under uncertainty, rather than on ethical arguments. Schneider, Traeger & Winkler (2012) extend equation (2) to account for overlapping generations. They reveal strong normative assumptions underlying the positive approach and show that the usual arguments of the normative approach are complicated by an equity trade-off between generations alive at the same point in time versus equity over generations across time.

## 2 Discounting under Uncertainty

The social discount rate relies on future consumption growth, which is uncertain. Within the standard model, only strongly serially correlated or catastrophic risk have a serious impact on the discount rate. We briefly discuss two extensions that incorporate general risk aversion and ambiguity aversion into the social discount rate, showing that these can have a major impact on the discount rate. We close the section commenting on Weitzman's (2009) Dismal Theorem and the Weitzman-Gollier puzzle.

### 2.1 Stochastic Ramsey equation

The social discount rate under uncertainty is generally defined using a certain consumption trade-off as in section 1.1 shifting consumption into an uncertain future. Then, the resulting consumption discount factor  $\Gamma_t$  captures the present value of a certain consumption unit in an uncertain future. As a consequence, the right hand side of equation (1), defining  $\Gamma_t$ , gains an expected value operator expressing that marginal utility gained from an additional future consumption unit is uncertain. For the subsequent analysis, we assume two periods, isoelastic preferences  $u(c) = \frac{c^{1-\eta}}{1-\eta}$ , and a normal distribution of the growth rate  $\tilde{g} = \tilde{c}_1/c_0$  with expected growth rate  $\mu$  and standard deviation  $\sigma$ . Then the consumption discount rate is

$$r = \delta + \eta\mu - \eta^2 \frac{\sigma^2}{2} . \quad (3)$$

The contributions of time preference and of expected growth coincide with the corresponding terms under certainty in equation (2). The third term  $-\eta^2 \frac{\sigma^2}{2}$  results from consumption growth risk and decreases the social discount rate, increasing the present value of a certain consumption unit in the future. It is proportional to the growth variance  $\sigma^2$  and the square of the consumption elasticity of marginal utility  $\eta$ . In the current context,  $\eta$  is frequently interpreted as a measure of risk aversion. However, it still is the measure of aversion to intertemporal substitution and section 2.2 explores a model incorporating general risk attitude.

We can interpret the timing in our setting in two different ways. First, the time between the first and the second period can be one year. Then,  $\delta$ ,  $\mu$ , and  $\sigma$  will generally be in the order of percent. For example, Kocherlakota (1996) estimates  $\mu = 1.8$  and  $\sigma = 3.6$  for almost a century of U.S. consumption data. Then, the risk term in equation (3) will be an order of magnitude smaller than the other terms: for  $\eta = 2$  ( $\eta = 1$ ) the growth contribution is  $\eta\mu = 3.6\%$  (1.8%) while the risk contribution is 0.5% (0.1%). Under the assumption of an iid growth process, equation (3) captures the constant, annual social discount rate.

Second, we can interpret period 0 as the investment time of a project, and period 1 as the payoff period. Assuming a constant annual expected growth rate, the two first terms on the right hand side of equation (3) increase linearly in time. Dividing the equation by the time span  $t$  between investment and payoff yields the average annual consumption discount rate. The first two contributions to this average rate are just as in (3), while the risk term becomes  $-\eta^2 \frac{\sigma^2}{2t}$ . For a random walk of the growth rate, the variance grows linearly in time, confirming the result that an iid growth process implies a constant annual discount rate. However, under serial correlation the variance increases more than linearly in time and the risk term increases the further the payoff lies in the future. Then, long term payoffs are discounted at a lower discount rate than short term payoffs: the term structure of the social discount rate decreases. Due to this finding, France and the U.K. adopted falling discounting schemes for project evaluation. We discuss the case of perfect serial correlation in more detail in section 2.5

## 2.2 General risk attitude

Equation (3) is based on the intertemporally additive expected utility model. In this model, the consumption elasticity of marginal utility has to capture two distinct preference characteristics: the propensity to smooth consumption over time and risk aversion. Positively, these two attitudes differ. Also normatively, there is no reason that the two should coincide.<sup>2</sup> In general, risk affects economic evaluation in two different ways. First, a stochastic process generates wiggles in the consumption paths. Agents dislike these wiggles if they have a propensity to smooth consumption over time. Second, agents dislike risk merely because it makes them uncertain about the future. This is an intrinsic aversion to risk that is not captured in the intertemporally expected utility standard model. The finance literature has successfully exploited general risk attitude to explain various asset pricing puzzles. In the context of determining the social discount rate, the most important puzzles solved are the equity premium and the risk free rate puzzles. Resolving these puzzles requires a model that gets the risk free rate right and explains the premium paid on risky equity. In a positive approach, where we measure preferences or interest rates based on market observations, it is important to use a model that gets these rates right. In a normative approach, the model forces the decision maker to think about both risk aversion and intertemporal (or intergenerational) consumption smoothing.

We keep the assumptions of a normal distribution of the growth rate and of isoelastic preferences, now with respect to both: consumption smoothing over risk and over time. Calling the measure of intrinsic risk aversion RIRA for relative intertemporal risk aversion, Traeger (2008) derives the social discount rate

$$r = \delta + \eta\mu - \eta^2 \frac{\sigma^2}{2} - \text{RIRA} \left| 1 - \eta^2 \right| \frac{\sigma^2}{2}. \quad (4)$$

Here, the parameter  $\eta$  only expresses the propensity to smooth consumption over time. The second term on the right hand side captures the growth

---

<sup>2</sup>It is a widespread misunderstanding that von Neumann & Morgenstern's (1944) axioms, underlying the expected utility model, together with additive separability on certain consumption paths, underlying the discounted utility model, would imply the discounted expected utility model. Merging these two assumptions results in a more general evaluation framework that distinguishes risk attitude from the propensity to smooth consumption over time (Traeger 2010b).



effect, while the third term captures the dislike of the agent for the wiggles in the consumption path generated by a stochastic process. The new term is proportional to the measure of intrinsic risk aversion, which is not captured in the standard model, and further reduces the discount rate. Increasing risk aversion (in the Arrow-Pratt as well as in the intrinsic sense) reduces the discount rate. In contrast, increasing  $\eta$  generally increases the discount rate. Disentangled estimates and calibrations in the asset pricing context result commonly in  $\eta = 2/3$  and  $\text{RRA} \in [8, 10]$  (Vissing-Jørgensen & Attanasio 2003, Basal & Yaron 2004, Basal, Kiku & Yaron 2010). Picking  $\text{RRA} = 9$  yields a coefficient of relative intertemporal risk aversion of  $\text{RIRA} = 25$  and a discounting effect of intrinsic risk aversion that is  $\text{RIRA} |1 - \eta^2| \frac{\sigma^2}{2} / \eta^2 \frac{\sigma^2}{2} \approx 31$  times larger than the effect created by aversion to consumption wiggles. In our numerical example with  $\mu = 1.8\%$  and a standard deviation of  $\sigma = 3.8\%$  the growth effect reduces to  $\eta\mu = 1.2\%$ , the standard risk to  $0.03\%$ , and the effect of intrinsic risk aversion equals  $0.9\%$ . Then, the social discount rate becomes  $\rho + 0.3\%$  and reduces almost to pure time preference, which the cited calibrations generally find to be close to zero as well. See Traeger (2008) for a sensitivity analysis and Gollier (2002) for a different treatment of discounting in the case of general risk attitude. Note that equations (3) and (4) hold not just for certain project payoffs, but also in the case where the project payoff is independent of baseline uncertainty. Traeger (2008) discusses the case of correlation between project payoff and baseline uncertainty.

### 2.3 General uncertainty attitude

In general, decision makers do not know the probability distributions governing the future with certainty. Situations whether the decision maker does not know the underlying probabilities are known as situations of ambiguity, hard uncertainty, or deep uncertainty (as opposed to risk). Klibanoff, Marinacci & Mukerji (2005) and Klibanoff, Marinacci & Mukerji (2009) develop a convenient model of decision making under ambiguity known as the smooth ambiguity model. It is similar to a standard Bayesian model, but distinguishes the attitude with respect to known probabilities (risk) from the attitude with respect to unknown probabilities (ambiguity), which are identified with the Bayesian prior. Traeger (2010a) generalizes the model and establishes its normative foundation: acknowledging the existence of different types of uncertainty, risk aversion measures depend on the type of

uncertainty a decision maker faces, even within a framework based on the von Neumann & Morgenstern (1944) axioms. The smooth ambiguity model corresponds to the special case where risk attitude coincides with the attitude for consumption smoothing, but differs from the attitude to ambiguity. The measure of ambiguity aversion is similar to that of intertemporal risk aversion; we denote the coefficient of relative ambiguity aversion by RAA. We assume once more isoelastic preferences and normal growth. However, this time the expected value  $\mu^*$  of the normal distribution is unknown. Given a particular value  $\mu^*$ , the standard deviation is once more denoted  $\sigma$ . The expected growth rate  $\mu^*$  is governed by a normal distribution with expected value  $\mu$  and standard deviation  $\tau$ . Traeger (2008) calculates the resulting extension of the Ramsey equation as

$$r = \delta + \eta\mu - \eta^2 \frac{\sigma^2 + \tau^2}{2} - \text{RAA} \left| 1 - \eta^2 \right| \frac{\tau^2}{2}. \quad (5)$$

The formula resembles equation (4) for intertemporal risk aversion. The difference are, first, that in the Bayesian model the overall uncertainty creating consumption wiggles is captured by the sum of the variance of both normal distributions. Second, intrinsic uncertainty aversion only affects second order uncertainty captured by  $\tau$ . Extending the model to disentangle risk aversion from both ambiguity aversion and the propensity to smooth consumption over time, implies that the Ramsey equation collects both terms, those proportional to intertemporal risk aversion in equation (4) and that proportional to ambiguity aversion in equation (5) (Traeger 2008). In the case of isoelastic preferences, intrinsic uncertainty aversion in the sense of intertemporal risk aversion and smooth ambiguity aversion always reduces the social discount rate. Gierlinger & Gollier (2008) and Traeger (2011a) establish general conditions under which general uncertainty lowers the social discount rate. The latter paper also shows how a decrease in confidence in the futurity of the growth forecast can lead to a falling term structure.

## 2.4 Catastrophic risk

Weitzman (2009) develops an argument that catastrophes would swamp the importance of discounting. In a Bayesian decision model with isoelastic preferences he assumes that the stochastic process governing growth is unknown.

In contrast to the analysis in sections 2.2 and 2.3, Weitzman does not incorporate general risk or uncertainty attitude. Instead of assuming a normal prior on expected growth, Weitzman puts an uninformative prior on the variance of the growth process. He shows that the resulting overall uncertainty is sufficiently fat tailed to imply an infinite consumption discount factor, implying an infinite weight on future consumption. Weitzman calls this result a Dismal Theorem. A simplified perspective on his result, neglecting the precise model of uncertainty and learning in Weitzman (2009), is that inserting enough uncertainty into the model implies that as  $\tau \rightarrow \infty$  in equation (5) the discount rate goes to minus infinity. In utility terms, the intuition for Weitzman's result is that his model exhibits a sufficiently slow decline of the probability mass characterizing that future consumption approaches zero and marginal utility infinity.<sup>3</sup> Weitzman makes the point that, even if we bound marginal utility away from minus infinity, the discount factor would be highly sensitive to the precise bound.

The social discount rate here and in Weitzman's calculation gives the value of a certain marginal consumption unit shifted into the future. Weitzman constructed an immensely uncertain future and then calculates the value of handing the future agents the first certain unit. If such a certain transfer mechanism would be available, this transfer should happen. With the first unit transferred infinity goes away and we can calculate the optimal amount that should be transferred into the future. The discount rate is like a price. If we offer an agent dying of thirst in the desert the first sip of water, he would likely give up all his worldly belongings in exchange. However, this measurement would not generally give us the market value of water. If, in contrast, uncertainty is insuperable, then we cannot simply calculate the social discount rate based on a certain consumption transfer, but have to account for uncertainty in the transfer and its correlation to baseline uncertainty (Traeger 2008). The empirical plausibility of the magnitude of uncertainty that Weitzman (2009) assumes is also questionable in the climate context in which it is motivated. See Millner (2011) for a discussion and extension of Weitzman's model.

---

<sup>3</sup>It might be useful to step back from elaborate economic welfare representations. In terms of preferences Weitzman's model contains a zero consumption state and, probabilistically, a lottery that is to be avoided by all means. Weitzman shows that the willingness to pay to get rid of this state are 'all means'. Note that the expected utility model with isoelastic utility does not satisfy the usual choice axioms when including the zero consumption level.

## 2.5 The Weitzman-Gollier puzzle

Weitzman (1998, 2001) and Gollier (2004) analyze the social discount rate in the presence of uncertainty about future economic productivity. Both authors assume perfectly serially correlated interest rates. Weitzman derives a falling term structure and Gollier derives an increasing term structure from assumptions that are apparently the same. This finding is known as the Weitzman-Gollier puzzle. Two insights help to resolve the puzzle. First, the original papers on the puzzle did not take into consideration the change of marginal utility over time and risk states (Gollier 2009, Gollier & Weitzman 2010, Freeman 2010). Second, Gollier's reasoning is concerned with the uncertain payoff of an investment project, while Weitzman's reasoning gets at growth uncertainty changing baseline consumption in the future. Gollier asks for the following certainty equivalent discount rate: what average annual productivity must a certain project have in order to match expected annual productivity of the uncertain project? The term structure of this rate generally increases: the payoffs of the uncertain project grow exponentially over time under full serial correlation, and the highest interest rate scenario dominates the (linear) expected value. In contrast, Weitzman's suggested rate is in the spirit of equation (3), which implies a falling term structure under serial correlation.<sup>4</sup> If the payoff uncertainty of the project under evaluation is independent of the market interest, then the value of the uncertain project increases over time with respect to a certain project as Gollier's discount rate implies. Both the certain and the uncertain project increase in value over time in a world of serially correlated uncertainty, relative to a world of certainty, as reflected by Weitzman's decreasing discount rate. In the case where project payoff and market interest are perfectly correlated the effect pointed out by Gollier vanishes. Then, the exponential payoff growth emphasizing the high payoff states is offset by the simultaneous decrease of the marginal utility obtained from an additional consumption unit, because the realization of the project payoff simultaneously determines the total consumption growth in the economy.

---

<sup>4</sup>Weitzman (1998, 2001) argues only by means of productivity in the economy. However, a close examination of his argument shows that the relation between consumption growth and productivity growth makes his formula almost correct Gollier & Weitzman (2010). It is only almost correct because it overlooks that the consumption share generally responds to the resolution of uncertainty over the market interest (Freeman 2010, Traeger 2012b).

### 3 Extensions

This section surveys some important extensions of the Ramsey formula. We start by relaxing the assumption of an aggregate consumption good, and analyze how limited substitutability between environmental services and produced consumption affects the discount rate. We then discuss the case of hyperbolic discounting as triggered by a non-constant RPTR. Finally, we explain that the explicit treatment of overlapping generations generally leads to a model equivalent to that of non-constant RPTP.

#### 3.1 Environmental versus produced consumption

Above we assumed the existence of an aggregate consumption commodity. This assumption becomes crucial if different classes of goods are not perfect substitutes. In particular, produced consumption is likely to be an imperfect substitute for environmental goods and services. Moreover, the provision and consumption of environmental goods and services does not generally grow at the rate of technological progress. Then, as our economy grows, environmental goods and services become relatively more valuable over time. We can incorporate this effect into a cost benefit analysis by introducing a time dependent conversion factor that translates the costs and benefits in terms of environmental good into current value produced consumption units. Alternatively, we can price environmental and produced consumption by present value prices and apply good-specific discount rates for cost benefit analysis. In both approaches, the underlying discount rate is affected by imperfect substitutability.

We assume that a representative agent derives utility  $u(c_t, e_t)e^{-\rho t}$  from consuming produced goods  $c_t$  and environmental consumption and services  $e_t$ . We define the discount factor of the consumption good as above and the discount factor for the environmental good as the amount of the environmental good that an agent is willing to give up in the present in order to receive an additional unit of the environmental good in the future. This rate is known as the ecological discount rate.<sup>5</sup> The discount rate characterizing the rate of

---

<sup>5</sup>This rate indicates how the value of a unit of the environmental good changes over time. If we are concerned how much of a consumption unit in the present an agent should give up for a future unit of environmental services, then we simply have to multiply the

change of the discount factor for consumption becomes

$$\delta_c(t) = \rho + \eta_{cc}(t) \hat{c}(t) + \eta_{ce}(t) \hat{e}(t) \quad (6)$$

with  $\hat{c}(t) = \dot{c}_t/c_t$ ,  $\eta_{cc}(t) = \eta_{cc}(c_t, e_t) = -\frac{\partial^2 u}{\partial c^2} c / \frac{\partial u}{\partial c}$  and  $\eta_{ce} = \eta_{ce}(c_t, e_t) = -\frac{\partial^2 u}{\partial c \partial e} e / \frac{\partial u}{\partial c}$ . Unless both goods are perfect substitutes ( $\eta_{ce} = 0$ ), the consumption discount rate for produced consumption depends both, the growth of produced consumption and on environmental growth (or decline).

Assuming Cobb-Douglas utility  $u(c_t, e_t) = c_t^{a_c} e_t^{a_e}$  (where  $a_c + a_e = 1$ ) eliminates the overall growth effect because Cobb-Douglas utility is linear homogeneous. We use this function form to focus on the effect of growth differences between produced and environmental consumption. Then, the consumption discount rate for the produced good (6) simplifies to

$$\delta_c(t) = \rho + a_e(\hat{c}_t - \hat{e}_t) .$$

Relatively faster growth in produced consumption increases the produced consumption discount rate. Similarly, this faster growth of produced consumption reduces the discount rate for the environmental goods and services:

$$\delta_e(t) = \rho - a_c(\hat{c}_t - \hat{e}_t) .$$

Thus, if produced consumption grows more rapidly than consumption of environmental goods, the discount rate to be applied in a cost benefit analysis for environmental good preservation is lower than the discount rate for produced consumption. This adjustment of the social discount rate for the environmental good reflects an increase in the relative scarcity of the environmental good causing its (relative) price to increase. For constant growth rates, both social discount rates are constant. However, this is a consequence of the unit elasticity of substitution between environmental and produced consumption. In general, these good-specific discount rates change over time. Both the discount rate for produced consumption as well as the discount rate for environmental goods and services can fall over time as a consequence of limited substitutability (Hoel & Sterner 2007, Traeger 2011b).

---

corresponding ecological discount factor with the relative price of the two goods in the present.

### 3.2 Hyperbolic discounting

Many models of dynamic public policy involve non-constant social discount rates. The nature of the resulting policy problem depends on whether this non-constancy causes time inconsistency. Time inconsistent policies can imply an ongoing revision of the formerly optimal policy, even in the absence of new information. In contrast, a declining term structure caused by falling growth rates, serially correlated uncertainty, or limited between-good substitutability leads to time consistent plans. Here we analyze the most common model giving rise to non-constant discount rates that cause time inconsistent plans: models employing a non-constant RPTP.

Ramsey (1928) noted "... My picture of the world is drawn in perspective. . . I apply my perspective not merely to space but also to time." The obvious meaning of "perspective applied to time" is that events in the more distant future carry less weight today, just as objects in the distance appear smaller. Any positive discount rate, including constant discounting, creates this type of perspective applied to time. However, perspective means more than the apparent shrinking of distant objects. The simplest model of perspective applied to space, known as "one point perspective", can be visualized as the appearance of railroad tracks viewed straight on, so that the two rails appear to converge at the horizon. The distance between adjacent railroad ties appears to grow smaller the more distant are the ties, but the rate of change appears to fall (Karp 2009). This kind of perspective means that not only do distant objects appear smaller, but also that we are less able to distinguish between the relative size of two objects, the further they are from us. Hyperbolic discounting, which assumes that the discount rate falls over time, is the time analog of this spatial perspective.

Hyperbolic discounting arises in both behavioral models of individual decision problems (Laibson 1997) and in long-lived environmental problems (Cropper, Ayded & Portney 1994). In the former setting, individuals' tendency to procrastinate is a prominent rationale for hyperbolic discounting. The rationale in the environmental setting is more closely tied to the fact that the problem of interest (e.g. climate change) occurs on a multi-generation scale. If we care less about our grandchildren than we do about our children, and care still less about generations that are more distant from us, our preferences are consistent with a positive discount rate on the generational time scale. If, in addition, we make less of a distinction between two contiguous

generations in the distant future compared to two generations close to us, our pure rate of time preference is hyperbolic. We might have a preference for our children relative to our grandchildren but scarcely distinguish between those born in the 20'th and the 21'st generation from ours. If individuals have this kind of time perspective and if the social planner aggregates the preferences of agents currently alive, then the social planner has hyperbolic discounting.

Non-constant discounting arising from preferences, as described above, causes optimal programs to be time inconsistent. That is, at any point in time the current social planner would like to deviate from the plan that was optimal for an earlier social planner. The time inconsistency is easiest to see using a discrete time example of the “ $\beta, \delta$  model”, where the sequence of discount factors used at  $t$  to weigh payoffs at times  $\tau \geq t$  is  $1, \beta, \beta\delta, \beta\delta^2, \beta\delta^3, \dots$ . If  $\beta = \delta$  the discount factor is constant, and discounting is exponential. If  $\beta < \delta$ , discounting is “quasi-hyperbolic”. Consider a project that reduces time  $t+1$  utility by  $\frac{\beta+\delta}{2}$  units and increases  $t+2$  utility by 1 unit, and suppose  $\beta < \delta$ . A planner at time  $t$  would accept this project, because the present value of the utility loss,  $\beta\frac{\beta+\delta}{2}$ , is less than the present value of the utility gain,  $\beta\delta$ . However, the planner at time  $t+1$  rejects the project, because for that person the present value of the utility loss is  $\frac{\beta+\delta}{2}$ , which is greater than the present value of the utility gain,  $\beta$ . The case  $\beta < \delta$  is associated with procrastination: a tradeoff that looks attractive when viewed from a distance becomes less attractive when viewed from close up. If a unit of time is taken to be the span of a generation, quasi-hyperbolic discounting implies that we are willing to make smaller sacrifices for our children than we would like them (and all subsequent generations) to make for their children.

One resolution to the time-inconsistency problem assumes that the initial planner chooses the current action under the belief that her entire sequence of preferred actions will be carried out. This resolution is dubious in a multi-generation context, where a current decision maker is unlikely to believe that she can set policy for future generations. A second resolution is to treat the policy problem as a sequential game amongst policymakers (Harris & Laibson 2001, Karp 2007). The optimal action for a social planner at time  $t$  depends on her belief about how policymakers will behave in the future. In a Markov perfect equilibrium, actions, and therefore beliefs about future actions, are conditioned only on directly payoff-relevant state variables. Often those variables have a physical interpretation, e.g. an environmental stock.



### 3.3 Overlapping generations

A closely related explanation for non-constant discounting rests on a model of overlapping generations. Suppose that agents discount their own future flow of utility at a constant pure rate of time preference,  $\rho$ , and that in addition they discount the welfare of the not-yet born at rate  $\lambda$ . Agents with both “paternalistic” and “pure” altruism care about the utility flows of future generations; for these agents,  $\lambda < \infty$ , and for agents without altruism,  $\lambda = \infty$ . Agents with pure altruism – unlike agents with paternalistic altruism – consider the effect on intermediate generations of the welfare of more distant generations (Ray 1987, Andreoni 1989).

If agents’ lifetimes are exponentially distributed, with no aggregate uncertainty, all agents currently alive have the same expected lifetime (Yaari 1965, Blanchard 1985). Absent other considerations (e.g. different levels of wealth, because older agents have had more opportunity to accumulate) agents currently alive are identical, so there is a representative agent in the usual sense. If instead, agents’ have random lifetimes with finite support (Calvo & Obstfeld 1988) or finite deterministic lifetimes (Schneider et al. 2012), the older agents have shorter remaining (expected) lifetimes. In this case, a social planner, perhaps a utilitarian, aggregates the preferences of agents alive at a point in time.

For the case of exponentially distributed lifetimes and paternalistic altruism, the discount factor of the representative agent is the weighted sum of two exponentials (Ekeland & Lazrak 2010). (Models with pure and paternalistic altruism are observationally equivalent.) The associated discount rate is non-constant, except for the two limiting cases,  $\lambda = \rho$  or  $\lambda = \infty$ ; in the first limiting case, the social discount rate is constant at  $\rho$  and in the second it is constant at  $\rho +$  the death rate. If  $\infty > \lambda > \rho$  the social discount rate increases over time asymptotically to  $\rho +$  the death rate. If  $\lambda < \rho$  the social discount rate decreases over time asymptotically to  $\lambda$ . For both  $\lambda < \rho$  and  $\rho < \lambda < \infty$  agents have a degree of altruism and non-constant discounting, but only  $\lambda < \rho$  corresponds to hyperbolic discounting. We noted above that many have argued that the only ethical choice is to treat all generations symmetrically, regardless of their date of birth. In this context, that requires  $\lambda = 0$ , so that the social planner’s evaluation of the stream of an agent’s utility does not depend on when she was born.

The previous section explains why a time inconsistency problem arises when

discount rates are non-constant. As noted above, one resolution is to consider a Markov perfect equilibrium in the game amongst generations. A second resolution eliminates the time inconsistency problem by assuming that the social planner at any point in time discounts the utility of those currently alive back to their time of birth, rather than to the current time.

## References

- Andreoni, J. (1989), ‘Giving with impure altruism: applications to charity and ricardian equivalence’, *Journal of Political Economy* **97**, 1447–1458.
- Basal, R., Kiku, D. & Yaron, A. (2010), ‘Long run risks, the macroeconomy, and asset prices’, *American Economic Review: Papers & Proceedings* **100**, 542–546.
- Basal, R. & Yaron, A. (2004), ‘Risks for the long run: A potential resolution of asset pricing puzzles’, *The Journal of Finance* **59**(4), 1481–509.
- Blanchard, O. J. (1985), ‘Debts, deficits and finite horizons’, *Journal of Political Economy* **93**, 223–247.
- Broome, J. (1992), *Counting the Cost of Global Warming*, White Horse Press, Cambridge.
- Calvo, G. & Obstfeld, M. (1988), ‘Optimal time-consistent fiscal policy with finite lifetimes’, *Econometrica* **56**, 411–32.
- Cropper, M. L., Ayded, S. K. & Portney, P. R. (1994), ‘Preferences for life-saving programs: How the public discounts time and age’, *Journal of Risk and Uncertainty* **8**, 243–266.
- Ekeland, I. & Lazrak, A. (2010), ‘The golden rule when preferences are time inconsistent’, *Mathematical and Financial Economics* **4**(1).
- Freeman, M. C. (2010), ‘Yes, we should discount the far-distant future at its lowest possible rate: A resolution of the weitzman-gollier puzzle’, *The Open-Access, Open-Assessment E-Journal* **4**.
- Gierlinger, J. & Gollier, C. (2008), ‘Socially efficient discounting under ambiguity aversion’, *Working Paper*.

- Gollier, C. (2002), ‘Discounting an uncertain future’, *Journal of Public Economics* **85**, 149–166.
- Gollier, C. (2004), ‘Maximizing the expected net future value as an alternative strategy to gamma discounting’, *Finance Research Letters* **1**(2), 85–89.
- Gollier, C. (2009), ‘Should we discount the far-distant future at its lowest possible rate?’, *The Open-Access, Open-Assessment E-Journal* **3**.
- Gollier, C. & Weitzman, M. L. (2010), ‘How should the distant future be discounted when discount rates are uncertain?’, *Economic Letters* **107**, 350–353.
- Harris, C. & Laibson, D. (2001), ‘Dynamic choices of hyperbolic consumers’, *Econometrica* **69**(5), 935–957.
- Harrod, S. R. F. (1948), *Towards a dynamic economics*, Macmillan, London.
- Hoel, M. & Sterner, T. (2007), ‘Discounting and relative prices’, *Climatic Change* **84**, 265–280.
- Karp, L. (2007), ‘Non-constant discounting in continuous time’, *Journal of Economic Theory* **132**, 557 – 568.
- Karp, L. (2009), ‘Sacrifice, discounting and climate policy: five questions’. Giannini Foundation working paper.
- Klibanoff, P., Marinacci, M. & Mukerji, S. (2005), ‘A smooth model of decision making under ambiguity’, *Econometrica* **73**(6), 1849–1892.
- Klibanoff, P., Marinacci, M. & Mukerji, S. (2009), ‘Recursive smooth ambiguity preferences’, *Journal of Economic Theory* **144**, 930–976.
- Kocherlakota, N. R. (1996), ‘The equity premium: It’s still a puzzle’, *Journal of Economic Literature* **34**, 42–71.
- Koopmans, T. C. (1963), On the concept of optimal economic growth, Cowles Foundation Discussion Papers 163, Cowles Foundation, Yale University.
- Laibson, D. (1997), ‘Golden eggs and hyperbolic discounting’, *The Quarterly Journal of Economics* **112**(2), 443–477.

- Millner, A. (2011), ‘On welfare frameworks and catastrophic climate risks’, Working Paper.
- Nordhaus, W. (2008), *A Question of Balance: Economic Modeling of Global Warming*, Yale University Press, New Haven. Online preprint: A Question of Balance: Weighing the Options on Global Warming Policies.
- Nordhaus, W. D. (2007), ‘A review of the Stern review on the economics of climate change’, *Journal of Economic Literature* **45**(3), 686–702.
- Pigou, A. C. (1932), *The economics of welfare*, 4th edn, Macmillan, London.
- Ramsey, F. P. (1928), ‘A mathematical theory of saving’, *The Economic Journal* **38**(152), 543–559.
- Ray, D. (1987), ‘Nonpaternalistic intergenerational altruism’, *Journal of Economic Theory* (41), 112–132.
- Schneider, M., Traeger, C. P. & Winkler, R. (2012), ‘Trading off generations: Infinitely lived agent versus olg’, *CESifo Working Paper No 3743*.
- Solow, R. M. (1974), ‘The economics of resources of the resources of economics’, *American Economic Review* **64**(2), 1–14.
- Stern, N., ed. (2007), *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge.
- Traeger, C. P. (2008), ‘Why uncertainty matters - discounting under intertemporal risk aversion and ambiguity’, *CUDARE Working Paper No 1092*.
- Traeger, C. P. (2010a), ‘Subjective risk, confidence, and ambiguity’, *CUDARE Working Paper No 1103*.
- Traeger, C. P. (2010b), ‘Wouldn’t it be nice to know whether Robinson is risk averse?’, *CUDARE Working Paper No. 1102*.
- Traeger, C. P. (2011a), ‘Discounting and confidence’, *CUDARE Working Paper No 1117*.
- Traeger, C. P. (2011b), ‘Sustainability, limited substitutability and non-constant social discount rates’, *Journal of Environmental Economics and Management* **62**(2), 215–228.

- Traeger, C. P. (2012*a*), ‘Once upon a time preference - how rationality and risk aversion change the rationale for discounting’, *CESifo Working Paper No 3793* .
- Traeger, C. P. (2012*b*), ‘What’s the rate? Disentangling the Weitzman and the Gollier effect’, *CUDARE Working Paper No 1121* .
- Vissing-Jørgensen, A. & Attanasio, O. P. (2003), ‘Stock-market participation, intertemporal substitution, and risk-aversion’, *The American Economic Review* **93**(2), 383–391.
- von Neumann, J. & Morgenstern, O. (1944), *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton.
- Weitzman, M. L. (1998), ‘Why the far-distant future should be discounted at its lowest possible rate’, *Journal of Environmental Economics and Management* **36**, 201–208.
- Weitzman, M. L. (2009), ‘On modeling and interpreting the economics of catastrophic climate change’, *The Review of Economics and Statistics* **91**(1), 1–19. 06.
- Yaari, M. E. (1965), ‘Uncertain lifetime, life insurance and the Theory of the Consumer’, *Review of Economic Studies* **32**, 137–150.