Discounting: A Review of the Basic Economics

Geoffrey Heal†

I review the justifications given for discounting future benefits relative to present, and distinguish between the pure rate of time preference, or utility discount rate, and the consumption discount rate, also sometimes known as the social rate of discount. I discuss when to choose one or the other, and how to choose a discount rate in a real-world project.

I. FUNDAMENTAL DILEMMAS OF DYNAMIC WELFARE ECONOMICS

Environmental economics provides a good entry point to a discussion of discounting, as intertemporal issues play a particularly central role: the time scale of many environmental processes is radically longer than conventional economic time scales. Global warming and loss of biodiversity provide perfect illustrations. Global warming may have its main impacts on human societies one hundred or more years hence, and likewise the costs of loss of species diversity in terms of simplification of ecosystems and loss of genetic variability are likely to be felt most strongly by generations quite remote from us. This is not to say that there will be no short-run impacts from these phenomena. There will be, but they are likely to be dwarfed by consequences that will become apparent only over very long periods. So to assess and evaluate properly these changes we need to look relatively far into the future—possibly a century at least. More generally, environmental assets such as watersheds, species diversity, rangelands, marine ecosystems, and climate regimes are assets that are in principle very long-lived. They have functioned as they do today for millennia, and if well managed will continue to do so equally far into the future. In this, they are rather different from the assets that humans construct and that we are used to valuing. These typically have life spans measured in years or decades. So to appreciate fully the contributions that environmental assets can make to human welfare we need a very long view. This does not sit easily with the economist's standard practice of discounting the future at a real rate of at least 3 percent or 4 percent.

† Paul Garret Professor of Public Policy and Corporate Responsibility, Columbia Business School, and Professor, School of International and Public Affairs, Columbia University, gmh1@columbia.edu, http://www.gsb.columbia.edu/faculty/gheal. This Article draws heavily on my entry *Intertemporal Welfare Economics and the Environment*, in Karl-Göran Mäler and Jeffrey R. Vincent, eds, 3 *The Handbook of Environmental Economics: Economywide and International Environmental Issues* 1106 (Elsevier 2005).

After all, if we discount at 3 percent per year, then $100 one hundred years from now is worth only $5 now; thus events so far into the future will be of little consequence in a cost-benefit analysis. In an obvious and intuitive sense, discounting seems to tip the scales against the future.[1] Is this indeed the case? What are the arguments for discounting over such a long period? What are the alternatives, if any?

The debate to which these questions give rise is not a new one. It runs right back to the origins of dynamic welfare economics. Frank Ramsey, who wrote a paper on optimal growth from which we can still learn today, commented that "discount[ing] later enjoyments ... is ethically indefensible and arises merely from the weakness of the imagination."[2] And his contemporary Roy Harrod was equally outspoken, remarking that discounting is "a polite expression for rapacity and the conquest of reason by passion."[3] They are saying, rather clearly, that the right discount rate is zero. More recently, *The Economist*, normally a repository of mainstream thinking on economics, was driven to remark: "There is ... something awkward about discounting benefits that arise a century hence. For, even at a modest discount rate, no investment will look worthwhile."[4] Yet Ramsey's pointed remark, though intuitively appealing, misses some deep technical points relating to the ranking of alternative consumption streams over time. These are important issues conceptually, and also complex and indeed treacherous from a technical perspective.

In working through these issues, we need to start by understanding our options for ranking intertemporal utility or consumption streams. Consider an economic model in which each generation lives for a finite number of periods and generations may overlap in time. Many people—though not all—would agree with the idea that we ought to give equal weight to the welfare levels at all points in time. It is after all difficult to make a really strong case for treating some generations better or worse than others. Of course if we are utilitarians, we may place less weight on a marginal increment of consumption of rich generations than of poor ones, but this is not an intertemporal judgment. It is an interpersonal one. It arises from diminishing mar-

---

1   See Geoffrey Heal, *Nature and the Marketplace: Capturing the Value of Ecosystem Services* 167–69 (Island Press 2000) (describing the effects of applying discounting to environmental values).

2   F.P. Ramsey, *A Mathematical Theory of Saving*, 38 Econ J 543, 543 (1928) (examining the issue of how much a country should save over time).

3   R.F. Harrod, *Towards a Dynamic Economics: Some Recent Developments of Economic Theory and Their Application to Policy* 40 (Macmillan 1948) (claiming that a government doing what is best for its subjects will not discount).

4   *What Price Posterity*, Economist 73 (Mar 23, 1991), cited in Geoffrey Heal, *Valuing the Future: Economic Theory and Sustainability* 58–59 (Columbia 1998) (explaining that many have questioned the wisdom of discounting).

ginal utility and not from differential treatment of generations. Suppose then that we want to treat all generations equally, and that we are utilitarians at least in the limited sense that we assume at each date t welfare is a concave increasing function of consumption, $u(c_t)$. This assumption embodies the classical one of diminishing marginal utility. We want to maximize the total welfare, the sum of welfare over all periods

$$\sum_{t=1}^{t=T} u(c_t).$$

This is easy when we have a finite number T of periods: we just give equal weight to the utility of each, as we are doing here. But economists have typically wanted to work with infinite horizons. This is mainly because of a reluctance to specify a date beyond which nothing matters, as one does when one chooses a terminal date T that is finite. Also influential is a concern about the impact of end effects, by which I mean sensitivity of the ranking of paths to the precise end date specified. So we tend to look instead at the sum of utilities from the present and forever on, the infinite sum

$$\sum_{t=1}^{t=\infty} u(c_t).$$

In many interesting and relevant cases, this sum will be infinite. In fact, it will be infinite for all consumption sequences on which utility levels are bounded away from zero, and on many others as well. This means that we cannot now represent choosing the best consumption sequence as finding the highest value of an objective function, which is the standard way of finding a best choice. Finding the best path is not in this case a conventional maximization problem. The only way of making sure that the infinite horizon sum is a finite number, so that this is a standard maximization problem, is to treat generations unequally, and in particular to give little weight to "most" of them, which of course is what discounting does. So there is apparently a practical reason for discounting. It is a way of ensuring that we have a well-defined preference order over the set of alternative infinitely long-lived consumption sequences between which we must choose, and that the best according to this ordering is the solution to a maximization problem.

Is it the only way to solve the problem or are there others? Ramsey was clearly aware of the difficulty back in the 1920s and had an ingenious alternative. His approach ensures that we give equal weight to all generations, but depends on some special assumptions and is not

a general resolution of the issues.[5] Others have developed more general approaches. For example, Carl Christian von Weizsäcker[6] and others have suggested the overtaking approach, which is now widely used. This ranks as best the consumption sequence, if any, whose cumulative utility sum eventually exceeds that on any other path. Formally this means that path c_t ranks above path c_t^1 if and only if there exists a time T^1 such that whenever $T > T^1$, then

$$\sum_{t=1}^{t=T} u(c_t) > \sum_{t=1}^{t=T} u(c_t^1).$$

This is a nice idea: we compare the cumulative welfare levels along alternative paths and see if there is one that eventually dominates. The idea behind this approach is that we can avoid discounting future utility levels and thus treat all generations equally. It is not completely successful in this.[7] And, once again, it does not allow the best choice to be found as the solution to a conventional maximization problem. But for most purposes it is a good solution.

The conclusion of this discussion is that we must be ingenious if we are to evaluate consumption sequences over time in a manner that gives equal weight to all generations, if at the same time we insist on working with infinite horizons and a roughly utilitarian framework. Perhaps in this case we should waive the requirement of infinite horizons? In fact, as we will see below, this concession buys surprisingly little.

A second problem, less fundamental but nonetheless demanding and only recently moving to the center of the stage, is dynamic consistency. A choice of a consumption path is dynamically consistent if it has the following property: if at some date during the execution of the chosen path we stop and ask what path we would now choose, given what we have done to date, then (provided that no parameters have changed) the answer is that we continue with the original choice. In other words, we see no reason to revise our choice merely because of the passage of time. Not all algorithms for making choices over time

---

5   Ramsey made some rather special assumptions that ensured that this sum converged in his case. He assumed that utility levels are bounded above and then sought to minimize the total shortfall over time of actual utility levels from their maximum level:

$$\sum_{t=1}^{t=\infty} [b - u(c_t)].$$

See Ramsey, 38 Econ J at 545 (cited in note 2). Here b is the upper bound of the utility function: think of it as "Bliss." Id.

6   See generally C.C. von Weizsäcker, *Lemmas for a Theory of Approximate Optimal Growth*, 34 Rev Econ Stud 143 (1967) (discussing the best utility path to take for preferred levels of economic growth).

7   See Heal, *Valuing the Future* at 65–67 (cited in note 4).

are dynamically consistent. Indeed, relatively few are. If we feel the need for dynamically consistent choices over time, then this again constrains how we can approach this area.

Finally, a third issue is that a path of consumption over time should generally be intertemporally efficient. This means that it is a path with the property that no variation will make some generation better off and none worse off. Obviously, if such a variation were possible, we would normally want to take advantage of it.

Finding a way to make choices that attain or come close to these three desiderata—equal treatment of all generations, consistent choices, and efficient choices—is difficult. It is the subject matter of dynamic welfare economics. As noted, it is particularly relevant in the environmental context because of the unusually long time horizons implied by the unfolding of anthropogenically induced modifications to the biosphere.

II. Discounting Utility or Consumption?

As a vehicle for illustrating the ideas surrounding discounting, consider an intertemporal optimization problem in which we seek to maximize the integral of welfare over an infinite horizon, where welfare depends on the flow of consumption and on a state variable (the state of the environment perhaps) and where there is a tradeoff between consumption and the state variable:

$$\text{Max} \int_0^\infty u\left(c_t, s_t\right) e^{-dt} \text{subject to} \quad c + \frac{ds}{dt} = f(s). \tag{1}$$

Here c is consumption of a final good, and s is a stock that could be a capital stock or an environmental state variable. This is a standard optimal growth model. Output $f(s)$ can be consumed or invested. We are discounting future utilities at the rate d. If $d > 0$, this means that future generations are being given less weight than the present generation just because they are in the future: futurity alone condemns them to a discount relative to us. A very important distinction in this type of utilitarian model is between the rate d at which *utility* is discounted, also known as the social rate of time preference or pure rate of time preference, and the rate at which *consumption* is discounted. The latter rate is sometimes called the social discount rate. The distinction is simple. Within this framework, one can ask the following question: Suppose the economy follows an optimal time path, one that solves the above problem, and we consider adding an increment of consumption at some date t. What is the value of this incre-

ment in terms of its contribution to the objective function, and how does this value change as the date t is changed?[8]

This contribution to the objective function is the value that should be assigned to an increment of consumption. The rate at which this contribution changes over time is the consumption rate of discount. It is the rate at which the weight of an increment of consumption changes over time. It is the rate at which incremental units of consumption are implicitly discounted along the path.

Clearly, the value of an increment of consumption c at date t is the marginal utility of consumption at t discounted to the present[9]

$$u_c(c_t,s_t)e^{-dt}c,$$

and the rate at which this value changes with t is

$$\frac{1}{u_c(c_t,s_t)e^{-dt}} \times \frac{du_c(c_t,s_t)e^{-dt}}{dt},$$

which we can easily compute to be

$$-d - \eta_{c,c}\frac{dc/dt}{c} - \eta_{c,s}\frac{ds/dt}{s}$$

where $\eta_{c,c} = -cu_{c,c}/u_c > 0$ is the elasticity of the marginal utility of consumption with respect to the level of consumption and $\eta_{c,s} = -su_{c,s}/u_c$ is the elasticity of the same quantity with respect to the level of the stock of the asset. These measure how quickly the marginal utility of consumption changes in response to changes in the levels of consumption and of the state variable. This expression is the consumption discount rate. Consider for simplicity the case in which the utility function is additively separable, so that the cross derivative is zero and $u_{c,s} = 0$. Then the consumption rate of discount is

$$d + \eta_{c,c}\frac{dc/dt}{c}.$$

---

8   For a mathematical explanation of the distinction between the utility discount rate and the consumption discount rate, see Geoffrey Heal, *Depletion and Discounting: A Classical Issue in the Economics of Exhaustible Resources*, in Robert W. McKelvey, ed, 32 *Proceedings of Symposia in Applied Mathematics: Environmental and Natural Resource Mathematics* 33, 39–40 (American Mathematical Society 1985).

9   u_c and u_s are the first partial derivatives of the function u with respect to its arguments c and s. $u_{c,c}$, etc., are likewise the second partial derivatives, using obvious notation.

For a linear utility function, or for variations in the level of consumption small enough that a linear approximation to the utility function suffices, this reduces to the utility rate of discount: the two concepts are the same. In this case, we have a positive consumption rate of discount if and only if we have a positive utility rate of discount. As Partha Dasgupta, Karl-Göran Mäler, and Scott Barrett note, the consumption discount rate is in general not constant over time, as the terms

$$\eta_{c,c} \frac{dc/dt}{c} - \eta_{c,s} \frac{ds/dt}{s}$$

are time varying.[10]

More generally, the two discount rates—the consumption discount rate and the utility discount rate—differ. And, if consumption falls over time so that

$$\frac{dc/dt}{c} < 0$$

then the consumption discount rate may be negative, so that the weight given to a increment of consumption actually rises over time (a point also emphasized by Dasgupta, Mäler, and Barrett). In fact, if an economy is following an optimal path in the utilitarian sense, then the first-order conditions for optimality give us further information about the consumption discount rate.

Intuitively, the consumption discount rate will differ from the utility discount rate because welfare levels change over time. If, for example, future generations are richer than the present generation, then within a utilitarian framework the value of a marginal unit of consumption to them will be less than to us, and this will be reflected in the consumption discount rate. This is the genesis of the term

$$\eta_{c,c} \frac{dc/dt}{c}$$

in the expression above, as this term shows how marginal utility is falling because consumption is rising. If consumption were to be falling rather than rising over time, this effect would go into reverse and future increments of consumption would be more highly valued than present ones. The discount rate could be negative.

One obvious and general proposition is the following. If the utility function has an elasticity of marginal utility of consumption that is

---

10   See Partha Dasgupta, Karl-Göran Mäler, and Scott Barrett, *Intergenerational Equity, Social Discount Rates, and Global Warming*, in Paul R. Portney and John P. Weyant, eds, *Discounting and Intergenerational Equity* 51, 53 (Resources for the Future 1999).

bounded, and if the utilitarian-optimal path tends in the limit to a stationary solution, then *at this stationary solution the consumption discount rate is equal to the utility discount rate.* Another way of saying this is that the social discount rate equals the pure rate of time preference. This is immediate from the expression

$$d + \eta_{c,c}\, \frac{dc/dt}{c}$$

for the consumption discount rate: if dc/dt is zero at a stationary solution, then this expression is just the utility discount rate. So, in general, the utility and consumption discount rates converge in the limit along utilitarian-optimal paths, or indeed any paths that have limits.

Consider some particular cases in more detail, firstly the pure depletion model of Harold Hotelling (this corresponds to the model (1) above when $f(s) = 0$ for all s). The first-order condition is that

$$\frac{dc/dt}{c} = -\frac{d}{\eta}$$

so that along an optimal path the consumption discount rate is always zero, whatever the utility discount rate.[11] In fact this is obvious. The first-order condition for optimality in the Hotelling problem is just that the marginal contribution of an increment of consumption to the objective should be the same at all times, which is precisely that the consumption discount rate be zero. On an optimal path in the Hotelling model, *whatever the utility discount rate, and however uneven the distribution of consumption between generations, the consumption discount rate is zero.* This shows that a zero consumption discount rate does not imply any degree of equality of consumption over time. Note that in this case there is no stationary solution to the first-order conditions for optimality so that the consumption and utility discount rates do not converge.

For the model specified above, the consumption rate of discount along a utilitarian optimal path is

$$\frac{u'_s}{u'_c} + f'$$

which is generally positive and again goes to the utility discount rate in the limit. In this case, the first order conditions admit a stationary

---

[11]   See Harold Hotelling, *The Economics of Exhaustible Resources*, 39 J Polit Econ 137, 139–40 (1931) (defining an "exhaustible resource" as an "absolutely irreplaceable" asset, that is, an asset that does not replenish at all over time).

solution so that the utility and consumption discount rates—pure rate of time preference and social rate of discount—converge in the limit. A similar convergence holds for models with capital accumulation and production.[12]

In summary, the consumption and utility discount rates, or social discount and time preference rates, are not independent concepts: when there is a stationary solution to the utilitarian problem, they are equal asymptotically along a utilitarian optimal path, and are always linked by the first-order conditions. Furthermore, the fact that the consumption discount rate is zero does not imply any degree of inter-generational equity, as the Hotelling example shows clearly. Having a zero consumption discount rate is not a solution to the ethical problem that led Ramsey and Harrod to decry the discounting of future utilities.

The central point is that we cannot avoid the need to choose a utility discount rate by focusing instead on consumption discount rates. Nor can we justify a positive utility discount rate by invoking an argument that the consumption discount rate may be quite different. The consumption discount rate is driven by the utility discount rate, the form of the utility function, and the technology of the economy.

Which of these two rates should be used for cost-benefit analysis? The answer should be clear from the analysis so far. The utility dis-count rate is a general equilibrium concept used in models of the evolution of an entire economy over time. If we have a planning model of the economy as a whole and are using this to assess how much to de-vote to preventing climate change, the general equilibrium concept is appropriate: we should use the utility discount rate. If, however, we are evaluating a small project that will have no economy-wide implica-tions—say the conservation of a regional forest or fish stock—then this is a partial equilibrium exercise, and the consumption discount rate is appropriate. I will return to this issue in the concluding section.

How much are the difficulties that arise in discussing discount rates artifacts of the use of an infinite horizon? Surprisingly, the answer is very little. Certainly one of the fundamental difficulties is treating all generations equally when the horizon is infinite. However, we have a partial solution to this, in the form of the overtaking criterion. This gives us an infinite-horizon utilitarian approach with a zero discount rate, that is, equality of generations.

Perhaps more important than this is the fact that problems with large finite horizons do not really look that different from infinite-horizon problems. By this, I mean that the solution to a finite-horizon

---

12   See Heal, *Valuing the Future* at 128–53 (cited in note 4) (exploring sustainability of both renewable and nonrenewable resources in economies with capital accumulation and production).

problem with a very long horizon looks for most of the long horizon
like the solution to the infinite-horizon problem. In fact, for suffi-
ciently long horizons the difference between the finite- and infinite-
horizon solutions is close to zero for much of the time. This idea lies
behind the Hammond-Mirrlees concept of an "agreeable plan,"[13] and
is discussed in Heal.[14] Working with infinite horizons is an analytical
convenience and puts some central issues into a sharp perspective, but
does not distort the issues. None of the discussion of consumption ver-
sus utility discount rates, for example, is dependent on the choice of
time horizon.

III. EMPIRICAL EVIDENCE

There is interesting evidence that in making choices over time
people use a framework different in certain salient respects from the
standard discounted utilitarian approach. Of course, even if we have a
clear picture of how individuals form their judgments about the rela-
tive weights of the present and future, this does not necessarily have
normative implications. We might still feel that relative to some appropri-
ate set of ethical standards they give too little (or too much) weight to the
future, and so are an imperfect guide to social policy.[15] However, in a de-
mocratic society, individual attitudes towards the present-future tradeoff
presumably have some informative value about the appropriate social
tradeoff and have at least an element of normative significance.

There is a growing body of empirical evidence suggesting that the
discount rate that people apply to future projects depends on, and
declines with, the futurity of the project.[16] Over relatively short periods
up to perhaps five years, people use discount rates that are higher
even than many commercial rates—in the region of 15 percent or in

---

[13] See generally Peter J. Hammond and James A. Mirrlees, *Agreeable Plans*, in James A.
Mirrlees and N.H. Stern, eds, *Models of Economic Growth: Proceedings of a Conference Held by
the International Economic Association at Jerusalem* 283 (Macmillan 1973).

[14] See Heal, *Valuing the Future* at 114 n 12 (cited in note 4) ("The zero discount rate utili-
tarian path that approaches the [highest environmentally sustainable utility level] is probably an
'agreeable plan,' in the sense of Hammond and Mirrlees.").

[15] See generally Geir B. Asheim, *Ethical Preferences in the Presence of Resource Con-
straints*, 23 Nordic J Polit Econ 55 (1996) (analyzing the suitability of different ethical preferences
in reaching ethically acceptable results in a capital accumulation and resource depletion model).

[16] See, for example, Maureen L. Cropper, Sema K. Aydede, and Paul R. Portney, *Prefer-
ences for Life Saving Programs: How the Public Discounts Time and Age*, 8 J Risk & Uncertainty
243, 251–55 (1994) (describing empirical research suggesting that the discount rates individuals
apply decrease as a function of time delay); George Loewenstein and Drazen Prelec, *Anomalies
in Intertemporal Choice: Evidence and an Interpretation*, in George Loewenstein and Jon Elster,
eds, *Choice Over Time* 119, 120–21 (Russell Sage 1992); George Loewenstein and Richard H.
Thaler, *Anomalies: Intertemporal Choice*, 3 J Econ Perspectives 181, 183–84 (1989); Richard
Thaler, *Some Empirical Evidence on Dynamic Inconsistency*, 8 Econ Letters 201, 205 (1981).

some cases substantially more. For projects extending about ten years, the implied discount rates are closer to standard rates—perhaps 10 percent. As the horizon extends, the implied discount rates drop to in the region of 5 percent for thirty to fifty years and down to of the order of 2 percent for one hundred years. The empirical evidence also indicates that the discount rate used by individuals, and the way in which it changes over time, depends on the magnitude of the change in income involved.

A.   Logarithmic Discounting and the Weber-Fechner Law

This empirically identified behavior has been termed "hyperbolic discounting"[17] and is consistent with a very general set of results from the natural sciences, which find that human responses to a change in a stimulus are nonlinear and are inversely proportional to the existing level of the stimulus.

As an example, the human response to a change in the intensity of a sound is inversely proportional to the initial sound level. The louder the sound initially, the less we respond to a given increase. The same is true of responses to an increase in light intensity. These are illustrations of the Weber-Fechner law, which is formalized by the statement that human response to a change in a stimulus is inversely proportional to the preexisting stimulus. In other words, for a given change, the response is less, the greater the existing value. In symbols,

$$\frac{dr}{ds} = \frac{K}{s}$$

or, integrating,

$$r = K \log s$$

where r is a response, s a stimulus and K a constant.

The empirical results on discounting cited above suggest that something similar is happening in human responses to changes in the futurity of an event: a given change in futurity (for example, postponement by one year) leads to a smaller response in terms of the decrease in weighting the further the event already is in the future. This is quite natural. Postponement by one year from next year to the

---

[17]   See Loewenstein and Prelec, *Anomalies in Intertemporal Choice* at 126–27 (cited in note 16) (noting that a hyperbolic discount rate matches empirical findings); George Ainslie and Nick Haslam, *Hyperbolic Discounting*, in Loewenstein and Elster, eds, *Choice Over Time* 57, 57–92 (cited in note 16) (arguing that only hyperbolic discount rates can explain the pervasiveness of self-defeating behavior in human beings). See also David Laibson, *Life-Cycle Consumption and Hyperbolic Discount Functions*, 42 Eur Econ Rev 861, 864–69 (1998) (using a hyperbolic discounting model to explore various types of economically irrational human behavior).

year after is clearly quite a different phenomenon from postponement from fifty to fifty-one years hence. The former obviously represents a major change; the latter, a small one. If we accept that the human reaction to postponement of a payoff or cost by a given period of time is indeed inversely proportional to its initial distance in the future, then this suggests that the Weber-Fechner law can be applied to responses to distance in time as well as to sound and light intensity. The result is that the discount rate is inversely proportional to distance into the future. Put differently, we react to proportional rather than absolute increases in the time distance. Denote the discount factor at date t by $\Delta(t)$, so that this represents the weight placed on benefits at date t relative to a weight of unity at time zero. In this case the discount rate $q(t)$, which is minus the rate of change of this weight over time,[18] is $q(t) = -(d\Delta(t)/dt) / \Delta(t)$. We can formalize the idea that a given increase in the number of years into the future has an impact on the weight given to the event, which is inversely proportional to the initial distance in the future as:

$$q(t) = -\frac{1}{\Delta} \times \frac{d\Delta}{dt} = -\frac{K}{t}$$

or

$$\Delta(t) = e^{-K \log t} = t^{-K}$$

for K a positive constant. Such a formulation has several attractive properties.[19]

A discount factor

$$\Delta(t) = e^{-K \log t}$$

has an interesting interpretation: the replacement of t by $\log(t)$ implies that we are measuring time differently, by equal proportional increments rather than by equal absolute increments. We react in the same way to a given percentage increase in the number of years hence of an event, rather than to a given absolute increase in its number of years hence. We shall call this "logarithmic discounting." This approach is quite consistent with that of acoustics, where, in response to the Weber-Fechner law, sound intensity is measured in decibels, which re-

---

[18]   The inclusion of a minus sign is required because it is conventional to report the discount rate as a positive number, whereas the rate of change of the discount factor is normally negative.

[19]   The discount rate q goes to zero in the limit, the discount factor (t) also goes to zero and the integral

$$\int_1^\infty \Delta(t)\,dt = \int_1^\infty e^{-K \log t}\,dt = \int_1^\infty t^{-K}\,dt$$

is finite for K positive and greater than unity. The lower limit of integration in this example is one, not zero, as t^{-K} is ill-defined for $t = 0$ and $K > 0$.

spond to the logarithm of the energy content of the sound waves, and not to energy content itself. In general, nonconstant discount rates can be interpreted as a nonlinear transformation of the time axis.

IV. CHOICE OF A DISCOUNT RATE

Economists are often asked — by international agencies, government officials, or members of environmental organizations — how in practice one should choose the discount rate to be used in project evaluation, or whether one should move away from discounting and use another approach. It should be clear from the discussion above that there are no simple answers to such questions. However, it is possible to give some general criteria that may be useful.

One can start with trying to clarify whether discounting utility or consumption is at stake. Do we need a pure rate of time preference (utility discount rate) or a consumption discount rate (social discount rate)? The former is appropriate when we are dealing with decisions that will affect the entire growth path of an economy or a region. To rephrase, utility discounting is appropriate when we are working with a general equilibrium model and general equilibrium consequences will follow from the choices under consideration. By contrast, discounting consumption is appropriate when we are working in a partial equilibrium context and the underlying growth path and resource allocation of the economy can be taken as given. In such situations we are considering changes that will amount to marginal alterations of the initial situation. Some examples can help to clarify. Alterations in economic policy designed to reduce greenhouse gas emissions in an industrial country are probably in the first category. They could be sufficiently far-reaching to alter the general equilibrium of the economy. So could a decision about the construction of a large dam in a small developing country. But with a purely local decision, such as the conservation of a local fishery or forest, it is clearly appropriate to view this in a partial equilibrium framework. We are considering marginal alterations about the economy's initial pattern of resource allocation.

As noted above, choosing utility and consumption discount rates involve different issues, and in general the former is a necessary but not sufficient condition for the latter: the consumption discount rate depends on, but is not fully determined by, the utility discount rate. This makes it appropriate to start with a discussion of the choice of the utility discount rate. This, of course, presumes the choice of a utilitarian framework, or one like it, by which I mean a Ramsey-esque or overtaking approach. All of these require the selection of either a single utility discount rate, which may be zero, or a schedule of time-varying discount rates as implied, for example, by the choice of a constant logarithmic discount rate. The only approach that lies outside

this general utilitarian-style framework is the Rawlsian, which I be-
lieve ultimately has less than the others to recommend it, particularly
in the intertemporal framework.[20]

Martin Weitzman recently conducted a survey of 2,160 profes-
sional economists, seeking their opinion on the appropriate choice of
discount rate for long-term environmental problems such as global
warming.[21] The modal rate recommended by this group was 2 percent,
the median 3 percent, and the mean 4 percent.[22] Unfortunately, it was
not clear from the survey whether the choice referred to a utility or a
consumption discount rate, although it is perhaps reasonable to as-
sume that respondents took the question to refer to utility discount
rates as there was no information about growth rates and the other
factors necessary to select a consumption discount rate.[23] These re-
sponses give us a clear indication that the majority of the economics
profession is aware of some of the issues discussed above. In general,
the choice of a discount rate for projects with a life of a decade or less
would be considerably above these rates. Clearly economists are se-
lecting lower rates in these responses in recognition of the impact of
"normal" rates over time horizons that are very long by conventional
economic standards. This does not prove that any particular answer or
approach is the correct one, but it does provide a degree of reassur-
ance that there is a general recognition of these problems. The positions
of Ramsey, Harrod, von Weizsäcker, and indeed of most economic
theorists and philosophers who have written on this, is that the utility
discount rate should be zero. Such a position is an ethical rather than an
economic judgment, and there is no obvious ethical reason why future
people should be considered less valuable than present people.

Whatever one's ultimate aim, the first move has to be a choice of
a long-term utility discount rate. We do not need to know risk premia
or the riskiness of any projects that may be undertaken. Nor do we

---

20   See Heal, *Valuing the Future* at 58–59, 76 (cited in note 4) (explaining the Rawlsian frame-
work and rejecting it in the intertemporal context because it does not allow sacrifice by the present
generation because it is likely to be the poorest generation in a world with economic growth); P.S.
Dasgupta and G.M. Heal, *Economic Theory and Exhaustible Resources* 270 (Cambridge 1979)
(arguing that we still do not know how individuals will choose when faced with uncertainty).

21   See Martin L. Weitzman, *Gamma Discounting*, 91 Am Econ Rev 260, 266–69 (2001). See
also generally Martin L. Weitzman, *Why the Far-Distant Future Should Be Discounted at Its
Lowest Possible Rate*, 36 J Envir Econ & Mgmt 201 (1998) (arguing that uncertainty about inter-
est rates in the distant future suggests that the lowest possible discount rate should be used in
cost-benefit analyses of long-term projects).

22   See Weitzman, *Gamma Discounting*, 91 Am Econ Rev at 268 (cited in note 21).

23   See id at 266 (reproducing the survey, which asked: "Taking all relevant considerations
into account, what real interest rate do you think should be used to discount over time the (ex-
pected) benefits and (expected) costs of projects being proposed to mitigate the possible effects
of global climate change?").

need to be aware of the nature of distortions in the economy and the impact of these on the relationship between market and shadow prices. Nor, finally, do we need to know the rate of return on capital investments or the cost of money. At an equilibrium or an optimum these are a function of the utility discount rate, and not the other way around. At issue are the relative weights to be placed on welfare levels occurring at different dates: should these weights decline with futurity and, if so, according to what pattern? This judgment about intertemporal distribution is at the heart of the choice of a utility discount rate. This judgment has to be made first. Given such a judgment, we can work out the consumption discount rate from the formulae above.

Recall that the consumption discount rate, or social rate of discount, is the utility discount rate plus terms that measure the rate of change of marginal utility. These terms depend on the curvature of the utility function—its elasticity—and on the rates of change of consumption:

$$d + \eta_{c,c} \frac{dc/dt}{c} + \eta_{c,s} \frac{ds/dt}{s}.$$

For a conventional Ramsey-type model without stock externalities (so that $\eta_{c,s} = 0$ and the third term is zero) the first-order conditions for optimality imply that

$$d + \eta_{c,c} \frac{dc/dt}{c} = f'(s)$$

where f' is the marginal product of capital s. This states that *the consumption discount rate*

$$d + \eta_{c,c} \frac{dc/dt}{c}$$

is equal to the marginal product of capital f'. Note that here both the left and right hand sides—social discount rate and rate of return—are endogenous to the solution of the model and are driven by the utility discount rate and the preferences and technology. This makes an important point, namely that in a general equilibrium framework we should not use the historical return on capital as a utility discount rate (in contrast with the positions of Weitzman[24] and William Nordhaus[25]).

---

24   See Martin L. Weitzman, *"Just Keep Discounting, But ..."*, in Portney and Weyant, eds, *Discounting and Intergenerational Equity* 23, 25 (cited in note 10) (recognizing that "long-term productivity has grown over time at a more or less trendless rate" and arguing that this will continue into the future).

25   See William D. Nordhaus, *Discounting and Public Policies That Affect the Distant Future*, in Portney and Weyant, eds, *Discounting and Intergenerational Equity* 145, 150 (cited in note 10) (using "observed returns on capital" when discounting).

In a partial equilibrium situation with no stock effects on welfare, however, we can take the social rate of discount to be the return on capital.

With a stock-dependent utility function, the social discount rate is

$$d + \eta_{c,c} \frac{dc/dt}{c} + \eta_{c,s} \frac{ds/dt}{s}$$

and from the first-order conditions[26] on an optimal path

$$d + \eta_{c,c} \frac{dc/dt}{c} = f_k .$$

Hence, in a model with both natural and produced capital the social rate of discount or utility discount rate is

$$f_k + \eta_{c,s} \frac{ds/dt}{s} .$$

For a separable utility function, $\eta_{c,s} = 0$, we again have equality between the utility discount rate and the return on capital, but not in general. In general, we expect that the elasticity of marginal utility with respect to the stock variable is not zero, as the utility of many goods will be affected by the state of the environment. If a better environment enhances the values of other goods—environmental stocks and other goods are complements—then

$$\eta_{c,s} = - su_{c,s} /u_c < 0$$

so that the utility or social discount rate is greater or less than the return on capital when the environmental stock is falling or rising respectively. If environmental stocks and other goods are substitutes, these inequalities are reversed. In such cases, we cannot compute the utility discount rate from the return on capital without information about preferences, about complementarities or substitutabilities between environmental stocks and other goods, and about movements in environmental stocks.

In summary, we are working either with a general equilibrium or a partial equilibrium framework. In the former case we need a utility discount rate. This reflects an ethical judgment and is not obtained from economic data such as the return on capital, the risk premium, etc. If we are in a partial equilibrium situation, we use the consumption or social discount rate. We can use the return on capital as a start-

---

[26] See Heal, *Valuing the Future* at 141–53 (cited in note 4).

ing point for such calculations. However, we have to modify this by the term

$$\eta_{c,s} \frac{ds/dt}{s},$$

and the net result could be zero or negative if the environmental stock is changing. The result is also time varying.

All of these comments on the consumption discount rate are premised on the assumption of a first-best economy, for example an economy with no distortions and with a fully optimal allocation of resources. To understand this point, and to begin to understand what would happen if we were to drop this assumption, recall that there is a standard duality between shadow prices arising from intertemporal optimization problems and competitive market prices. The shadow prices could also emerge as market-clearing prices in a competitive economy with a complete set of markets. A complete set of markets in the present context means a complete set of futures markets, from the present to the infinite future. Given such a set of markets, a competitive economy would attain an intertemporal equilibrium at which prices were identical to the shadow prices and would follow the path described by the conditions for optimality. An alternative to the assumption of a complete set of futures markets would be the assumption of fully rational expectations. This amounts to the same thing. The only difference is in the packaging. Under these admittedly rather strenuous assumptions, we can think of f_k as the market return on capital and then talk about how to adjust this to calculate the social rate of discount to be used in a market economy.

If the assumption of a fully first-best economy is not met, it becomes much harder to characterize the factors that determine the consumption rate of discount. The assumption of a fully first-best economy may fail in several ways. One, as noted, is the lack of futures markets, or, equivalently, imperfections in capital markets. There is an equivalence here because futures markets are used for moving consumption and income over time, and that is also what capital markets do. We can therefore think of futures markets as devices for borrowing and lending. Another possible source of departure from first-best is the presence of wedges between borrowing and lending rates: this could be a consequence of taxes on income or of capital rationing—another aspect of capital market imperfections. In addition, there are reasons for differences from first-best that are more specifically related to the environmental nature of the problems under consideration. For example, the environmental stocks considered in the model above are often public goods—forests, stocks of biodiversity, climate regimes—and of course the investor in these will often have difficulty

in appropriating the returns that the investment generates for society as a whole. Likewise, there may be external effects driving a wedge between the private and social returns to investment.

Given differences between borrowing and lending rates, between private and social returns, and other deviations from the first-best framework implicitly assumed above, what is the right choice of a consumption discount rate? Unfortunately, this is an extremely complex subject, and one with no easy generalizations. Ideally we would modify the model to reflect the precise departures from first-best that are relevant in a particular situation, and then use this revised model to derive conclusions about the consumption discount rate, using the same mathematical methods as above. As a crude illustration of the model above, suppose that capital market imperfections made it impossible for agents in this model to save or dissave at more than a specific rate, say $\varepsilon > 0$. This would then impose a constraint that the difference between output and consumption (which is savings) must be less than this amount, so that in graphical terms the system would be forced to stay within an ε-neighborhood of the curve $c = f(s)$. This constraint would, in turn, have a shadow price that would interact with other shadow prices and affect the consumption rate of discount.

In practice there are too many different possible departures from first-best for it to be practical to model each particular case. Some general points, however, are obvious. We should try to correct returns for the differences between private and social costs, and should impute to investments in public goods the full social benefits resulting.

There is one less-obvious general point that is robust. Above, we noted that in a first-best situation with no stock externalities the consumption discount rate would equal the return on investment. The reason is that a small increase in consumption will lead to a small (equal) decrease in investment, and the returns to the two must be equal on the margin on an optimal path. Suppose that there is a wedge between these two returns: which is the more appropriate as a discount rate, if either? In the presence of taxes on income, the return on investment is typically greater than the consumption rate of discount. The former is before tax and the latter after tax. In considering the rate at which the valuation of consumption changes over time, we should note that a change in the time pattern of consumption will alter the time path of investment and of capital and output, which will have further implications for future consumption. A reduction in future consumption will make society immediately worse off, but may make society better off by increasing output and consumption at a later date. With a first-best allocation of resources all of this is reflected in shadow or market prices. With a second-best, it is not. The rate of change of the marginal utility of consumption, which is the

consumption rate of discount and is also the rate of change of the shadow price of investment, does not reflect changes in the level and valuation of output if there is a difference between the consumption discount rate and the return on investment.

Some commentators have suggested that the appropriate response is to use a weighted average of the rate of change of the marginal utility of consumption and the return on capital as a consumption discount rate.[27] A better approach is to assess the impact of consumption changes on the level of investment and to value these changes in investment by the shadow price of capital. This shadow price of capital at date t reflects the full contribution that extra capital available at date t makes to consumption from t onwards.[28] We then value the sequence of consumption and investment changes resulting from a change in policy in terms of consumption (the shadow price of capital makes this conversion for changes in capital) and discount them to the present at the consumption rate of discount. We are using the consumption rate of discount as a discount rate, but the return on investment is being used in calculating the shadow prices of capital. Robert Lind provides an excellent introductory survey of these issues,[29] and David Bradford's paper is a readable and authoritative original source.[30]

---

27 See, for example, Robert H. Haveman, *The Opportunity Cost of Displaced Private Spending and the Social Discount Rate*, 5 Water Resources Rsrch 947, 950–51 (1969).

28 See Kenneth J. Arrow, *The Rate of Discount on Public Investments with Imperfect Capital Markets*, in Robert C. Lind, et al, *Discounting for Time and Risk in Energy Policy* 115, 118 (Resources for the Future 1982) (outlining how to mathematically predict consumer behavior for discounting purposes); Robert C. Lind, *A Primer on the Major Issues Relating to the Discount Rate for Evaluating National Energy Options*, in Robert C. Lind, et al, *Discounting for Time and Risk in Energy Policy* 21, 23 (using the concept of the shadow price of private capital to account for the opportunity cost of financing public investments); David F. Bradford, *Constraints on Government Investment Opportunities and the Choice of Discount Rate*, 65 Am Econ Rev 887, 889 (1975) (explaining consumption patterns in relation to government investments); Kenneth J. Arrow and Mordecai Kurz, *Public Investment, the Rate of Return and Optimal Fiscal Policy* 117–18 (Resources for the Future 1970) (discussing the hypothesis that consumption is a function of wealth and the interest rate); Kenneth J. Arrow, *Discounting and Public Investment Criteria*, in Allen V. Kneese and Stephen C. Smith, eds, *Water Research: Economic Analysis, Water Management, Evaluation Problems, Water Reallocation, Political and Administrative Problems, Hydrology and Engineering, Research Programs and Needs* 13, 17–19 (Resources for the Future 1966) (same).

29 See generally Lind, *A Primer on the Major Issues* (cited in note 28).

30 See generally Bradford, 65 Am Econ Rev 887 (cited in note 28).