Climate policies deserve a negative discount rate

Marc Fleurbaey

Princeton University

mfleurba@princeton.edu

Stéphane Zuber

CNRS-CERSES and Paris Descartes

stephane.zuber@parisdescartes.fr

Introduction

Since the Stern Review (Stern 2007) and the debate it has sparked off, the discount rate has been at the center of heated discussions about climate policies. In the very long run, the discount rate makes a huge difference in the evaluation of policies. The following table (Table 1) shows the minimum return that a $1 investment for the future should have to be considered better than consuming it now, depending on the discount rate that is adopted and depending on the horizon. The 1.4% discount rate is advocated by the Stern Review, but later Stern suggested that 2.7% might be a better figure. The table shows that this hesitation is not innocuous. Obviously, adopting a much higher discount rate as recommended by Nordhaus (around 5.5%) has even more extreme consequences.

Horizon	1.4%	2.7%	Ratio
50	2.00	3.79	1.89
100	4.02	14.36	3.57
200	16.13	206.11	12.78
500	1,044.67	609,848.27	583.77
1000	1,091,327.24	371,914,916,666.52	340,791.38

Table 1: The implications of different discount rates

Legend: With a 1.40% discount rate, a $1 investment today must yield at least $4.02 in 100 years; with a 2.70% discount rate, the number jumps to $14.36, which is 3.57 greater.

The thesis defended in this paper is that climate policies may justify the use of a negative discount rate for their evaluation. There are two important steps in the argument, each of which is an interesting separate thesis: 1) Different policies may be evaluated with different discount rates depending on what populations are impacted; 2) In the long run only the worst scenario for the worst-off fraction of the population counts.

Our thesis is at odds with the conclusion of Chapter 7 in Posner and Weissbach (2010 –hereafter PW), even though we share the same premises: impartiality between generations, compatibility with ethical principles, and taking opportunity costs into account. They advocate using the market interest rate as the discount rate for the selection of particular projects. It is an interesting question to understand how our similar premises can deliver very different practical conclusions. The main difference is that we disagree on how to make use of the discount rate. For us, it is a tool to assess and compare different consumption paths or money flows in terms of net present value; for PW, it serves to take into account the opportunity cost of the investment.

Our thesis is somewhat closer to Weitzman's and Gollier's arguments in favor of using a small, possibly a negative (Weitzman 1998, Gollier 2002), but involves different reasons. We will also discuss Weitzman's (2009) more extreme thesis about the possible justifiability of a "dismal", very negative discount rate.

The paper is structured as follows. In the next section we briefly review the arguments of the advocates of the descriptive and the prescriptive approaches, in the debate about discounting (the "ethicists" and the "positivists", as called by PW), and discuss agreements and disagreements with PW. Then in the next section we explain what we retain from Weitzman's line of argument, and what we don't retain. The following explains our core arguments. For the sake of an easy presentation, the bulk of the paper is formulated in the context of utilitarian reasoning, but we explain in the penultimate section why the utilitarian approach must be replaced with a more promising approach and how this can affect the discount rate. The final section concludes.

The descriptive-prescriptive debate

The opposition between a descriptive approach and the prescriptive approach is hard to understand when it is labeled in this way, as suggested by Arrow et al. (1996). It is equally puzzling when the "ethicists"-"positivists" labels are used. As PW write, "in the end, of course, the positivists' approach is worth nothing unless it can be defended on ethical grounds." (p. 150) So, the debate is not between ethics and something else, it is a debate within ethics.

The descriptive approach invokes two ethical arguments. The first is that market rates reflect the preferences of the population, so that it is undemocratic to propose using different rates (Nordhaus 2007). The climate economists who propose using lower rates for climate policies are imposing their views on a population that appears to care less about the future than they do. This first argument is unacceptable but the reasons why it cannot be accepted are far from simple.

There are several mistakes in the argument. The first is to seek to impose the population preferences on every evaluator. Obviously, there are many views in the population. If an evaluator wants to examine a development path with a great concern for the future, there must be some people in the population who share this concern. Even if nobody shared this concern, the evaluator might be right against everyone else. Just as there is freedom of thought, just as different political parties can have their own platforms, there should be a space for economic evaluation that embodies various views about social welfare and the principles of intergenerational equity.

The reply to this objection will certainly be: Any evaluation is admissible, but the government, in its decisions, cannot impose idiosyncratic views on the whole population. This is a powerful argument, even though history contains praiseworthy examples of governments imposing policies against the majority opinion (e.g., the abolition of death penalty in France in 1981). But this powerful democratic argument does not imply that the market interest rate should serve as the discount rate. It only requires a democratic debate to take place. This debate will have to ponder the various arguments underlying the computation of the discount rate. One cannot pretend to know the conclusions of this debate in order to prevent some propositions from reaching the debate. Democratic principles cannot be used to bar some (minimally sound) ethical principles from the forum.

One could still try to argue that the market does tell us something about the population preferences about the future. The market interest rates are determined by the joint effect of technical possibilities (the productivity of capital) and the willingness of investors to transfer wealth into the future, with a benefit. Just like the relative market price of oranges and pears implies that all buyers active on both markets are willing to trade oranges for pears at this relative price, the interest rate indicates the investors' and savers' common marginal rate of substitution across time. This is true but investors and savers make decisions to transfer wealth for themselves by a few years. If they were asked to transfer wealth to other people living all over the world in many decades, they might express very different preferences. The financial markets don't ask them this outlandish question, and therefore we cannot pretend to use their answer to a different question for this purpose. Observe, moreover, that the market interest rate also depends on the distribution of wealth in the population, which has no reason to be particularly democratic.

Even if there were markets in which people could express such preferences (private donations to environmental NGOs focused on the climate might be the relevant source of information), it is doubtful that such preferences would be more respectable than the outcome of an outright democratic debate involving the relevant expertise and considering the best ethical arguments.

In conclusion, if experts like Stern propose a series of reasonable arguments leading to the conclusion that climate policies should be *evaluated* with a discount rate that is much lower than the market interest rate, they cannot be dismissed as undemocratic and off track. They should be admitted to the democratic debate and their arguments should be carefully listened to.

The second argument used by the advocates of the market rate is that this rate measures the opportunity cost of resources. This is the main argument considered, and endorsed, in PW. It is this argument that leads PW to propose using the market interest rate. This argument is crystal clear. Suppose a climate policy costs $1 today and brings benefits worth $14.40 in one hundred years. According to Table 1, this policy is better than the consumption of the $1 today if the discount rate is 2.7% or lower. The objection is that investing the same amount at the market rate, which is supposedly greater, would bring greater benefits to the future. Using a lower discount rate than the market is therefore branded as a recipe for choosing dominated policies which either cost more today or pay less tomorrow, or both.

This argument is very simple and extremely powerful. But it aims at the wrong target. More precisely, it relies in our view on a misunderstanding of the role of the discount rate. The purpose of the discount rate is to make consumption levels or monetary values comparable across time. It makes it possible to compute the net present value (NPV) of any change to the status quo. If the NPV is positive, the change is an improvement. But this does not mean that this particular change is optimal. In order to choose the best policy or project, one must compare the NPV (computed with whatever discount rate seems appropriate) of *all* options, including ordinary market investments. Clearly, with this methodology, if one option costs less today or pays more tomorrow (or both) than another option, it will be deemed preferable, whatever the discount rate!

There is therefore no danger that adopting a lower discount rate than the market rate may induce inefficient (i.e., dominated at each period) choices. It will only imply making different choices among the efficient (i.e., undominated) options. With a lower discount rate, one will choose to invest more for the future, but one will never be tempted to invest at a low rate of return when a high rate of return is possible. If a business-as-usual investment policy that puts all savings in the financial market brings more benefits to the future generations than a mitigation policy aimed at curbing GHG emissions, even the most devoted disciple of the Stern Review will approve it.

To illustrate this point, Table 2 presents an example of four policies, with their undiscounted benefits and their NPV according to two different values of the discount rate (1.4% is from the Stern Review, 5.5% has been advocated by Nordhaus). Policy A is a market investment at 5.5%, policy B is a climate policy with a 3% return, policy C is a market investment at 2% (a lower rate is available for a longer horizon), policy D is a climate policy with a 2.5% return. Policy B is dominated by the market (policy A), and this is clearly identified by the net present value at any discount rate. In contrast, policy D is not dominated by the market. But should we prefer policy A or policy D? Referring to the market rate only serves to check if a policy is dominated by the market. One needs a discount rate in order to compute a sensible present value. The example shows that the difference between a discount rate of 1.4% and a rate of 5.5% is important. With the latter, policies C and D have negligible returns in present value and are almost equivalent to throwing money away. (Their net present value is slightly above -1.00, but very close to it.) With the former, in contrast, policy D is preferable to policy A.

Horizon	**A**	**B**	**C**	**D**
0	-1	-1	-1	-1
100	211	19	0	0
500	0	0	19957	230109
Present value				
1,40%	51,66	3,79	18,10	219,27
5,50%	0,00	-0,91	-1,00	-1,00

Table 2: How to choose policies with various discount rates

Legend: Policy A costs $1 today and pays $211 in 100 years;

at a 1.4% discount rate, the present value is $51.66;

at a 5.5% discount rate, it is $0.

In conclusion to this point, when PW write: “even if the ethicists’ arguments are entirely correct, we must still carefully consider the opportunity cost of projects and pick those with the highest returns,” we fully agree, and every reasonable “ethicist” should agree, too. But this does not imply that the market rate of return should be used for the evaluation of projects.

Note that the use of discount rates would be superfluous if the problem were to choose between policies with similar time profiles like policies A and B in Table 2, because their own rates of return can be directly compared. It is only when there are time trade-offs that computing the NPV becomes useful, as in the comparison between policies A and D. (Actually, as we will show later, projects with the same time profile may affect different populations, thus deserving different discount rates, which is equivalent to incorporating social benefits in the computation of their rate of return.)

There is another related methodological issue on which we disagree with PW. They propose to use the discount rate in a limited way: "Discounting… should be seen only as a method of choosing projects, not as a method of determining our obligations to the future." (p. 168) This separation is a direct consequence of the tension produced by their idea that one should use the market discount rate for the choice of projects, but nevertheless follow the ethicists to decide how much to save for the benefit of the future generations. In other words, two discount rates would be used in the methodology proposed by PW, although they do not make it fully explicit. The low discount rate of the ethicists would serve to check if more should be invested, whereas the choice of particular projects would use the market rate in order to make sure to pick efficient options.

There is no need for such a dichotomous methodology. The "ethically right" discount rate can be used both for the selection of projects and for deciding when to stop saving for the future, which constitute in fact one and the same set of decisions –selecting the projects includes choosing the amount that is invested. One will start with the highest-NPV investment plans (which are those with the highest rate of return for their particular time profile) and go on as long as the NPV of the remaining projects is positive. Note that the discount rate itself goes up in the process, because as more is invested for the future, the future generations grow better off, which tends to raise the discount rate (see below for an explanation of this phenomenon). Therefore a low discount rate advocated now, on the background of a business-as-usual scenario in which the future generations are in jeopardy, need not be the indication of the market rate that will prevail after the recommended investment has been done. The convergence value of the market rate will be somewhere between the initial market rate and the initial discount rate.

Finally, let us briefly consider another objection raised by PW against the ethicists. They claim that choosing projects as the ethicists propose may be futile when private decisions to dissave may partly undo the public investments. Again, the ethicists can only agree and proclaim their innocence. Their criteria are meant to bear on final consequences, not on mistaken estimates of the consequences. If a certain Ricardian equivalence implies that the government cannot influence the macroeconomic savings rate, there is no point for the government to try to change it and no point in applying any discount rate to this kind of decision. If public savings partly crowd out private savings, this must be taken into account.

In this section we have defended the ethicists and their prescriptive approach against the attacks of the positivists with their descriptive approach, but we haven't said a word, almost, about what the prescriptive approach says and how it computes its discount rate. To this we now turn.

Dismal ?

As announced in the introduction, for simplicity we will provisionally adopt the utilitarian way of defining social welfare, and more precisely assume that social welfare can be computed as the sum of $U(c_i)$, where c_i is the consumption level of individual i. What is important about this approach is that the function U is assumed to be the same for all individuals, which means in particular that there is no preference for earlier generations against future generations, and we consider that this function embodies the preferences of the evaluator about inequalities in consumption. We will even adopt an assumption that is common in the economic literature, and gives a special form to the utility function:

$$U(c) = \frac{1}{1-\rho}\left(c^{1-\rho} - \hat{c}^{1-\rho}\right),$$

where ρ can be interpreted as a coefficient of aversion for consumption inequality, and $\hat{c}$ is the minimum level of consumption that is required to make utility positive. With this function, the marginal utility is equal to $c^{-\rho}$, which makes things quite simple.

For the sake of simplicity we ignore the risk that future generations will not exist. This issue will be introduced later in the paper.

Suppose that c_i is reduced by a small amount $-\Delta c_i$ and c_j, which occurs t periods later, is increased by a small amount Δc_j. Is this good for social welfare? If the changes are infinitesimal, one can use the marginal utilities to evaluate the changes, and the variation in social welfare is then equal to

$$-U'(c_i)\Delta c_i + U'(c_j)\Delta c_j.$$

This can be expressed in present value by dividing every term by the marginal utility of c_i,

$$-\Delta c_i + \frac{U'(c_j)}{U'(c_i)}\Delta c_j,$$

and by comparing this expression with the discounted sum $-\Delta c_i + \frac{1}{(1+\delta)^t}\Delta c_j$, one obtains the discount rate by the formula

$$1 + \delta = \left(\frac{U'(c_j)}{U'(c_i)}\right)^{-1/t}.$$

The discount rate is a direct expression of the relative priority of the two individuals (or generations), modulated by the time distance between the two individuals. If the future individual is better off than the present individual, the expression is greater than one, i.e., the discount rate is positive.

This methodology gives the discount rate that can serve to evaluate small projects. Any project that yields a rate of return greater than the discount rate is beneficial to social welfare. For big projects, the marginal utilities are no longer acceptable in the computation and one has to make a direct evaluation of the change in social welfare.

When the marginal utility is equal to $c^{-\rho}$, the formula for the discount rate simplifies into

$$1+\delta=\left(1+g_{ij}\right)^{\rho},$$

where $g_{ij}=\left(\frac{c_j}{c_i}\right)^{1/t}-1$ is the annual growth rate of consumption between c_i and c_j. When g_{ij} is small, this formula can be approximated by the famous Ramsey formula $\delta=\rho g_{ij}$.

A reasonable value for inequality aversion is $\rho=2$ (the Stern Review took $\rho=1$) while a standard estimate for the growth rate is 1.3 (as in the Stern Review), which implies a discount rate of approximately 2.6% (Stern adds a 0.1% term due to the risk of extinction of humanity, we ignore this term for the moment). This discount rate may be much lower than the market rate of return, but, as explained in the previous section, this is not particularly threatening for this methodology and there is no danger of choosing dominated investment plans.

Weitzman has proposed two interesting arguments in favor of adopting even lower discount rates for investments that pay in the very long run. We propose to accept one and reject the other.

Let us start with the acceptable argument, first put forth in Weitzman (1998). We actually present a variant of it which enables us to connect the argument to Ramsey's formula, as in Gollier (2002). Suppose that there is uncertainty about future consumption, and that our criterion is the expected value of social welfare (which is also, in the case of utilitarianism, the sum of expected utilities). Let us again consider two small changes Δc_i and Δc_j. Unlike the level of consumption, these changes are certain. The change in social welfare is now equal to

$$-U'(c_i)\Delta c_i+EU'\left(c_j\right)\Delta c_j,$$

where E denotes the expected value, and the formula for the discount rate becomes:

$$1+\delta^*=\left(\frac{EU'\left(c_j\right)}{U'(c_i)}\right)^{-1/t}.$$

(δ^* now denotes the discount rate, and we keep the notation δ for the discount rate in a particular state of nature.) What Weitzman noticed is that this formula involves neither the expected value of

$1+\delta$, nor the expected value of $(1+\delta)^t$, but the expected value of $(1+\delta)^{-t}$. More precisely, one has

$$1+\delta^* = (E[(1+\delta)^{-t}])^{-1/t}.$$

Now, what is remarkable about this expression is that it has the form of a well-known quasi-arithmetic mean of the form $(E[x^{-t}])^{-1/t}$, which is known to converge to the minimum value of x when t tends to infinity. Therefore, in the very long run, the discount rate under risk converges to the lowest possible value of the risk-free discount rate. This is a remarkable result. For long term evaluations, one can focus on the worst scenarios in which future consumption is the lowest and the corresponding discount rate is the lowest.

PW propose an intuitive explanation of this reasoning directly based on Weitzman's formulation. For every possible value of c_j there is a corresponding discount rate δ. The expected present value of the investment, when all possible discount rates are considered, is then equal to

$$E\left[\Delta c_i + \frac{1}{(1+\delta)^t}\Delta c_j\right] = \Delta c_i + E\left[\frac{1}{(1+\delta)^t}\right]\Delta c_j,$$

and by comparing this expression to $-\Delta c_i + \frac{1}{(1+\delta^*)^t}\Delta c_j$, one directly obtains the desired formula.

This explanation remains, however, a little mysterious because it is not obvious why one should compute the expected present value of a project rather than some other formula. In fact, in general the expected present value is not correct. What should be computed is the ratio of the expected values of the marginal utilities,

$$\Delta c_i + \frac{EU'(c_j)}{EU'(c_i)}\Delta c_j,$$

but when there is no risk about the present consumption the expected value of the marginal social value of c_i is just the sure value $EU'(c_i) = U'(c_i)$, and therefore the ratio of expected values is, in this specific case, the expected value of the ratio.

To illustrate the result, consider the situation in which with 80% chance, the growth rate will be 1.3% on average in the future, but there is 20% chance that it will be zero. Let us retain $\rho = 2$, so that the risk-free discount rate is either 2.6% or 0%. This example shows that the convergence to the lowest value may be rather slow, but also that the discount rate is very quickly well below the average discount rate 0.8x2.6 = 2.08%.

Horizon	**Discount rate**

50	1,74%
100	1,35%
200	0,80%
500	0,32%
1000	0,16%

Table 3: Discount rate for different horizons

when the growth rate is either 2.6% or 0%

Weitzman (2009) proposed an even more striking ("dismal") thesis by arguing that the discount rate can be made arbitrarily close to -100% even in the short run. A discount rate at -100% means that the future has absolute priority and that we should, at least at the margin, transfer resources to the future at all costs.

A simplified version of the argument is the following. Suppose that the utility function takes the form $U(c) = \frac{1}{1-\rho}(c^{1-\rho} - \hat{c}^{1-\rho})$, that the distribution of future consumption is bounded below by $\hat{c}$, but that there is a positive (possibly small) probability that $c = \hat{c}$. The lowest risk-free discount rate is then given by the expression

$$1 + \delta = \left(\frac{U'(\hat{c})}{U'(c_i)}\right)^{-1/t} = \left(\frac{\hat{c}}{c_i}\right)^{\rho/t}.$$

This expression can be rendered as small as one wants, which makes δ tend to -100%, by letting $\hat{c}$ tend to zero. When the lowest possible risk-free discount rate tends to -100%, and its probability is bounded away from zero, the correct discount rate δ^* also tends to -100%, even in the short run. This is because the quasi-arithmetic mean $(E[x^{-t}])^{-1/t}$ tends to zero when one of the possible values of x tends to zero and $t > 0$.

This is not exactly the form taken by Weitzman's (2009) argument. He assumes that future consumption is given by an expression of the form $\alpha e^{-\beta T}$, where T is future temperature. This expression tends to zero only if the temperature tends to infinity, and the distribution of temperatures must therefore be unbounded if one wants to have a probability of consumption being arbitrarily low. In addition, he needs to assume that the distribution of high temperatures has a "fat tail" (which means that the probabilities of high temperatures are considerably larger than a Gaussian law would indicate) in order to obtain that the expression $EU'(c_j)$ tends to infinity when $\hat{c}$ tends to zero (i.e., the probability of having $c = \hat{c}$ must not decline too quickly when $\hat{c}$ tends to zero), so as to be able to conclude that the discount rate must then tend to -100%.

One unpalatable aspect of this fat-tail story is that it relies on an unrealistic formula for future consumption. One does not need to have infinite temperatures to see life plummet on Earth. The temperature of boiling water is more than enough, and much lower temperatures would in fact be sufficient to entail a complete disruption of life systems.

As we have seen, the simpler form of the argument presented earlier only needs to assume that there is a positive probability, however small but bounded from zero, that life will be endangered in order to obtain the dismal conclusion. This assumption is realistic, and Weitzman himself claims that one may assign as much as 1% to the chance of having temperature increases of more than 20°C.

Is the dismal conclusion therefore convincing? We do not think so, actually. The dismal conclusion needs to identify a situation in which a future catastrophe has positive probability and must be avoided at all costs. Is there such a situation? If we consider individual decision-making for an analogy, we may think that avoiding death is indeed something that individuals would be willing to sacrifice all their resources to avert. Therefore, similarly, if the evolution of the climate put the survival of our species at risk and there are ways to reduce this risk, shouldn't we devote all our resources to it?

The analogy with individual decision-making is helpful, but it does *not* prove what is said in the previous paragraph. Death cannot be avoided with the available technology, it can only be postponed. It may seem that individuals would be willing to sacrifice all their resources if Death knocked at their door in the morning, but we also observe that they are not willing to spend all their resources above subsistence to improve their survival curve. Individuals routinely take additional risks of death in the near or, even more frequently, in the distant future, in order to make life more enjoyable. There is an appearance of paradox here, but the explanation is simple. Let the individual utility per period of life be determined by the function $U(c) = \frac{1}{1-\rho}(c^{1-\rho} - \hat{c}^{1-\rho})$. Let the probability of $c = \hat{c}$ in all the next periods (which is equivalent to death) be p. How much is the individual willing to pay in the present period in order to put that probability to zero? To fix ideas, suppose there is only one more period to live, and the consumption level in that last period would be $c_0 > \hat{c}$. The individual's willingness to pay w is determined by the equality

$$(c^{1-\rho} - \hat{c}^{1-\rho}) + (1-p)({c_0}^{1-\rho} - \hat{c}^{1-\rho}) = ((c-w)^{1-\rho} - \hat{c}^{1-\rho}) + ({c_0}^{1-\rho} - \hat{c}^{1-\rho}),$$

which implies

$$c^{1-\rho} - (c-w)^{1-\rho} = p({c_0}^{1-\rho} - \hat{c}^{1-\rho}),$$

an expression which clearly shows that w is small when p itself is small. In fact the right hand side measures the utility loss of early death, and it is a finite amount that justifies a limited willingness to pay to avert it. The loss is greater when there are more future periods, and it would tend to infinity if the number of periods were to grow unbounded. Death becomes an infinite loss only when perpetual life is a possibility.

But then, why do we have the intuition that people would sacrifice everything superfluous to avert imminent death? This is actually what the above formula shows. If the certainty of an additional period of life has a certain value $p({c_0}^{1-\rho} - \hat{c}^{1-\rho})$, we are willing to sacrifice exactly this value from our current utility in order to obtain this guarantee. What we are willing to sacrifice measures what appears superfluous in the circumstances, so that it is tautologically true that we are willing to relinquish everything superfluous to prolong life. But we are not willing to sacrifice a lot to eliminate a small risk of death, or to put ourselves in such dire straits that the current or the prolonged life would not be worth living.

At the level of the human species, similar considerations apply, although there is the additional complication that preferences about the longevity of the species depend on ethical principles. In existing theories about the cosmos, however, the survival of the species cannot be perpetual, so that making the extinction occur earlier is a finite loss for any reasonable ethical theory, and against a finite loss we should be willing to sacrifice only a finite amount of resources. Therefore there is no reason to give absolute priority to the future.

Note, however, that the above dismal argument was about the discount rate, therefore only about marginal changes. Could it be that we should give absolute priority to the future but only for a small amount of effort, after which the priority becomes finite? As a matter of fact, even an absolute priority for the future *in the margin* lacks plausibility. The expression $1 + \delta = (\hat{c}/c_i)^{\rho/t}$ yields a finite value for every positive $\hat{c}$, therefore there is no particular reason to consider the case in which $\hat{c}$ tends to zero. Weitzman is right that this parameter then acquires special importance, but there is nothing surprising or dismal about the fact that the marginal utility of consumption in the worst possible case should be scrutinized with care and should play an important role in precautionary decisions.

Priority for the poor in the long run

We do agree with Weitzman, however, that a negative discount rate may be justified for climate policies. There is another line of argument that, combined with the phenomenon, described in the previous section, of convergence toward to the lowest rate in the long run, reinforces the presumption that negative values are relevant. We borrow this argument from Fleurbaey and Zuber (2012).

The debate about "the" discount rate is somewhat misleading because there is not a single discount rate but as many discount rates as there are distributions of costs and benefits among different populations. We have already seen this phenomenon when the discount rate to be used depends on the time lag between generations, as in Table 3.

More generally, the formula that determines the discount rate is about changes in the consumption of two individuals i, j. The value of the discount rate depends on the consumption levels of these two individuals. Imagine now that two individuals j, k from a future generation, not just one, will benefit from a change in their consumption, so that we have to deal with the formula

$$-U'(c_i)\Delta c_i + U'(c_j)\Delta c_j + U'(c_k)\Delta c_k.$$

As before, one can compute the present value by dividing by the marginal utility of c_i, and compare this with a formula involving person-to-person discount rates:

$$-\Delta c_i + \frac{U'(c_j)}{U'(c_i)}\Delta c_j + \frac{U'(c_k)}{U'(c_i)}\Delta c_k = -\Delta c_i + \frac{1}{(1+\delta_{ij})^t}\Delta c_j + \frac{1}{(1+\delta_{ik})^t}\Delta c_k.$$

Now this formula is structurally similar to the formula obtained in the case of risk. Imagine that j and k share the benefit of the investment in fixed proportions: $\Delta c_j = \alpha_j B$ and $\Delta c_k = \alpha_k B$. The above expression can then be written as

$$-\Delta c_i + \frac{1}{(1+\delta)^t}B,$$

for

$$\frac{1}{(1+\delta)^t} = \alpha_j \frac{1}{(1+\delta_{ij})^t} + \alpha_k \frac{1}{(1+\delta_{ik})^t}.$$

The same argument as in the previous section implies that in the very long run, i.e., when t tends to infinity, δ will converge to the smallest value of person-to-person discount rates. The smallest value is obtained for the individuals who are the worst off in the future generation. Therefore, in the very long run, only the worst off of the future generations matter. More precisely, one must focus on the worst off among those who benefit from the investment. Those whose share is null play no role in the formula.

A complication is that the investment cost is generally paid by several members of the present generation. If one thinks of a public policy such as a mitigation effort to reduce GHG emissions, many individuals may be involved. Let us therefore consider the problem when several people from the present generation contribute in fixed shares: $\Delta c_i = \alpha_i C$ and $\Delta c_l = \alpha_l C$. But to simplify the presentation, let us come back to a situation in which only one individual from a future generation stands to benefit from the investment. The formula is now the following:

$$-U'(c_i)\Delta c_i - U'(c_l)\Delta c_l + U'(c_j)\Delta c_j = -(U'(c_i)\alpha_i + U'(c_l)\alpha_l)C + U'(c_j)\Delta c_j.$$

The present value can therefore be written as

$$-C + \frac{U'(c_j)}{(U'(c_i)\alpha_i + U'(c_l)\alpha_l)}\Delta c_j = -C + \frac{1}{(1+\delta)^t}\Delta c_j,$$

for

$$(1+\delta)^t = \alpha_i(1+\delta_{ij})^t + \alpha_l(1+\delta_{lj})^t.$$

This formula has the opposite behavior as the previous one. When t tends to infinity, δ tends to the greatest value of the person-to-person discount rates. What is remarkable is that the greatest value is obtained for the worst off of the present generation, among those who share in the cost.

When many individuals share the cost now and many individuals share the benefit in the future, these two results remain jointly valid, even though the formula is more complicated: In the very long run, *the discount rate converges to the worst-off-to-worst-off discount rate*, among the individuals who are affected by the change in consumption to be evaluated. This holds whatever the shares, although, of course, the speed of convergence is influenced by the shares.

Table 4 illustrates this phenomenon with four individuals, two from each generation. Shares in cost and benefit are supposed to be equal (half-half) in every generation. In the present generation the poor has a consumption of 1 unit, the rich has a consumption of 5 units. The dynasty of the poor has a consumption growing at 1.3% per year, the dynasty of the rich enjoys a growth rate of 1.5%. We keep $\rho = 2$.

Horizon	**Discount rate**	**Poor-Poor**	**Rich-Poor**	**Poor-Rich**	**Rich-Rich**
50	2,63%	2,62%	-3,78%	9,87%	3,02%
100	2,63%	2,62%	-0,63%	6,39%	3,02%
200	2,63%	2,62%	0,98%	4,69%	3,02%
500	2,62%	2,62%	1,96%	3,69%	3,02%

1000	2,62%	2,62%	2,29%	3,35%	3,02%

Table 4: The discount rate at different horizons, for society as a whole (column 2) and the person-to-person discount rates (columns 3-6)

The table shows that the poor-to-poor discount rate is a good indication of the social discount rate in this context, even at a moderate horizon (but this depends on the shares and is not true in general). Observe that the rich-to-poor and the poor-to-rich discount rates change with the horizon because their relative consumption does not evolve according to a constant growth rate. In the beginning the poor dynasty remains poorer than the rich of the first generation, which justifies a negative discount rate, whereas the rich dynasty is much richer than the poor of the first generation, which justifies a high discount rate. In the long run, the relative consumptions tend to follow the growth rate of each dynasty, which explains the convergence toward a discount rate that is specific to the beneficiary rather than specific to the donor.

We now come to the main thesis of this paper. Why should climate policies be evaluated with a negative discount rate? The person-to-person discount rate is negative when the present donor is richer than the future beneficiary. Given the above result, if we consider long-run policies, the discount rate should be negative when the poorest contributors to the policy are richer than the poorest beneficiaries. It is plausible that many climate policies satisfy this condition. Mitigation efforts, when they are well conceived, should put the burden on the high emitters who are typically among the affluent members of the present generation, but they will benefit many members of future generations. Moreover, it is often said that the most vulnerable to climate change are the poorest, so that many beneficiaries in the future will be among the poorest of their generation. Can we hope that the poorest of future generations will be better off than the middle class of the present generation? This appears, sadly, unlikely. Therefore climate policies that avoid imposing a burden on the poor members of the present generation deserve to be evaluated with a negative discount rate.

Another element reinforces the thesis. Weitzman's result of a convergence toward the lowest discount rate in the case of risk combines with the result presented in this section. In the very long run, the discount rate converges to the worst-off-to-worst-off discount rate *of the worst-case scenario*. Therefore, even if there are favorable scenarios in which the destitute populations catch up and reach good standards of living, it is enough to assign a positive probability to dark scenarios in which the standards of living of the poorest stagnate in order to validate our conclusion about the negative discount rate for climate, especially mitigation, policies.

Of course, this does not mean that such policies should have greater priority than other policies such as redistribution toward the poor members of the present generation (Schelling 1995). The choice of the best policies, as we have seen in the second section, involves a comparison of present values, not just checking that the chosen policy improves on the status quo. At least, however, we want to argue strongly against the popular thesis that the market rate should be applied indiscriminately to the evaluation of all policies, independently of the affected populations.

Beyond utilitarianism

So far we have adopted the utilitarian approach, which indeed dominates in the debate about discounting for the long run. The utilitarian approach is quite acceptable in the absence of risk, because the utility function can then be chosen, as suggested above, to embody the aversion to inequalities in consumption that the evaluator endorses. Formally, the utilitarian social welfare can then also be adopted by prioritarians and egalitarians who accept the property of subgroup separability that underlies the additive form of the criterion. (Subgroup separability means that the evaluation of a change affecting a subgroup of the population can ignore the consumption level of the unconcerned individuals and focus on the affected subgroup only. By an important theorem due to Gorman and Debreu, subgroup separability implies that the evaluation criterion can be represented by an additive function.)

In the presence of risk, things are less easy. The coefficient of inequality aversion also becomes a coefficient of risk aversion if the utilitarian criterion is then applied as the sum of expected utilities (or equivalently, the expected sum of utilities). There is therefore a dilemma. Either one respects the risk aversion of the population (assuming away a potential heterogeneity of risk preferences across individuals), which severely constrains the degree of inequality aversion, or one adopts a coefficient of inequality aversion on the basis of ethical principles and then potentially imposes on the population a degree of risk aversion that appears paternalist.

Fleurbaey (2010) proposed a compromise. The idea is that respecting preferences is much less compelling under uncertainty than in a risk-free context, because in the context of risk, by definition individuals are not perfectly informed about the consequences of their decisions. In particular, respecting preferences under risk may even appear to betray informed preferences when the evaluator has information about the final distribution. Suppose for instance that individuals are willing to take a risk but that it is known in advance that the only consequence of this risk is a widening of inequalities, without any overall gain. At the individual level the risk may appear attractive, but at the social level it is already known that many will be unlucky and that they actually act against their true interests when they are willing to take the risk. When making a decision under risk, each individual focuses on his own payoffs and ignores the correlation with other individuals. A social evaluator can take account of this correlation and forecast how many individuals will turn out to have acted against their ultimate interests.

This observation leads to the conclusion that respecting risk preferences is not always necessary, but it also suggests that respecting risk preferences remains an attractive idea when there is perfect correlation between individuals, because in such a situation an evaluator cannot forecast if some of them are acting against their interests. Fleurbaey (2010) shows that when the requirement to respect risk preferences is limited to the case of perfect correlation, other criteria than utilitarianism become acceptable, permitting a greater degree of inequality aversion. There is a theorem stipulating that, under minimal conditions of rationality under uncertainty, all such criteria must take the form of the expected value of the equally distributed equivalent (EDE) utility, which is the level of utility that would yield the same social welfare if it were equally distributed across all individuals.

Let us illustrate this with a particular functional form. Suppose that $Eu(c)$ represents the risk preferences of the individuals, assuming away any heterogeneity across individuals in order to keep things simple and in line with the literature on discounting. Suppose that in absence of risk one would like to use the prioritarian criterion $\sum_i \varphi\big(u(c_i)\big)$. Then the social criterion may take the form

$$E\varphi^{-1}\left(\frac{1}{n}\sum_i \varphi\big(u(c_i)\big)\right),$$

where φ^{-1} denotes the inverse function and n is the number of individuals.

In Fleurbaey and Zuber (2012), we study how this kind of criterion can be used in the computation of the discount rate. What is important is that the discount rate can then be approximated by the usual discount rate obtained for the additive social welfare function $\frac{1}{n}\sum_i \varphi\big(u(c_i)\big)$, where n is the size of the population, to which one has to add a term that depends (positively) on the correlation between the well-being of the beneficiaries of the investment and social welfare at the global level. (There may be an additional term depending on the attitude of the criterion to population size – this issue will be explained later.)

It is not easy to figure out whether this result pushes in the direction of raising or lowering the discount rate for climate policies. A first issue is whether climate risks generate common risks for most populations or induce negative correlations. In the case of common risks, the correlation term is positive and tends to raise the discount rate. The case of negative correlations is possible if a change in the climate would actually be beneficial in the high latitudes where the most affluent populations are now settled, whereas it would be dramatic for the subtropical areas in which the most vulnerable populations live.

But even if negative correlations occur, it is still possible for the correlation term to be positive. Indeed, recall that in the long run, the poor members of the future generations are those who matter for the discount rate. If the degree of inequality aversion (i.e., the concavity of function φ) is strong, social welfare as measured by the EDE is then close to the lowest utility in society, and therefore directly correlated with the well-being of the worst-off.

Not much is known about the size of the correlation term and simulations are not easy to perform because they require considering scenarios that describe the situation of the whole human species, from beginning to end.

This compels us to mention another issue that cannot be ignored when the risk of extinction is considered. In the Stern Review there is a 0.1% additional term that comes from the estimated exogenous 1/1000 risk per annum of extinction of the species due to cosmic phenomena (meteors, eruptions) or unforeseen disruptions of life systems (pandemics). The underlying utilitarian reasoning is that the expected value of total utility is equal to

$$U(c_0) + \frac{999}{1000} U(c_1) + \left(\frac{999}{1000}\right)^2 U(c_2) + \cdots$$

The computation of the discount rate then involves an additional term in the formula of the marginal change in social welfare:

$$-U'(c_i)\Delta c_i + \left(\frac{999}{1000}\right)^t U'(c_j)\Delta c_j,$$

implying

$$1 + \delta = \frac{1000}{999} \left(\frac{U'(c_j)}{U'(c_i)}\right)^{-1/t},$$

which is approximately equivalent to adding 0.001 to the initial value (when the latter is close to 1).

There is, however, an issue that such an approach raises. Different values for the longevity of the human species imply different sizes for the total human population, which requires taking a stance on the question of the optimal size of the population. In the utilitarian galaxy, there are three popular approaches. Total utilitarianism, implicitly adopted above, adds utilities considering that a new member with a positive utility always improves social welfare. Critical-level utilitarianism adds utilities but deducts a fixed amount for every new member. In other words, it computes the sum of $U(c_i) - \alpha$, which means that adding a new member to society is considered beneficial only if his utility is above α. The introduction of the critical level, however, does not affect the marginal utility of consumption for existing members and therefore does not affect the discount rate. The third approach is average utilitarianism, which divides total utility by the size of the population and considers that adding new members is desirable only when their utility is above average. The computation of the discount rate for this third approach is substantially different and, to the best of our knowledge, has not been explored.

With the EDE criterion introduced in this section, one has various options for the critical level, but unlike utilitarianism there is only one constant critical level that can be taken, and this is the lowest possible utility. Another salient option is to take a critical level that is equal to the EDE itself. For a significant degree of inequality aversion, the EDE is close to the lowest utility in the population, which may not be an unreasonable option for the critical level. It is this option that we have retained in the above discussion of the correlation term. If one retained a critical level equal to the lowest possible utility (instead of the EDE), then one would have to add another term to the discount rate, that depends on the correlation between the well-being of the beneficiaries of the investment and the population size (see Fleurbaey and Zuber's paper for details).

To conclude this section, we provide a simple illustration of the discount rate obtained by the various utilitarian criteria and by the EDE in a simple two-state scenario. We assume that in the favorable state the human species spans 2 million years, comprising 80000 generations (a generation is 25 years), and we assume that 4000 generations (100,000 years) have already lived up to now. The world population is assumed stable from now on, with 3 billion members per generation (three generations overlap at any given moment in time). The past population since the origins is assumed to have grown from an initial number of two individuals at the growth rate of 0.3% per generation until 10,000 years ago, when the growth rate rose to 2.62%.

There are two dynasties, one consuming 1 unit today and the other consuming 5 units today. The evolution of consumption over time can hardly be assumed to be exponential over such a long horizon. Indeed, assuming that the first generation consumed one hundredth of current consumption (per capita), the growth rate of per capita consumption up to now would have been 0.115% on average per generation, which seems very small, but would imply that the last generation in 1.9 million years should consume 10^{38} as much as now, a big number that is probably much greater than the number of planets in the whole universe (the number of stars is estimated below 10^{24}).

We will instead assume a linear growth of consumption, with a constant inequality between the two dynasties. This still leaves the last generation with about 19 times our current consumption level. (In fact, our results do not depend much on either choice for consumption growth, because we focus on a short horizon of 40 generations, over which consumption growth is small in both descriptions.)

In order to introduce risk we also assume that with a 20% probability, consumption will stagnate forever from now on, and that only 40,000 generations live (one million years). This is the unfavorable state.

We retain the utility function $U(c) = \frac{1}{1-\rho}(c^{1-\rho} - \hat{c}^{1-\rho})$, and take $\varphi(u) = \frac{1}{1-\varepsilon}u^{1-\varepsilon}$, with $\rho = \varepsilon = 2$ and $\hat{c} = 1/200$. (This is half the first generation's consumption. Unlike in Weitzman's dismal result, ours do not depend much on the value of $\hat{c}$, except in the dismal example provided in Table 6 below.) We consider a policy that is paid only by the rich dynasty of the 4000th generation and benefits equally to every dynasty in future generations, i.e., we ask how much $1 equally shared between a rich and a poor in the future is worth compared to $1 paid by a rich today. The horizons in Table 5 are expressed in years, as in the previous tables.

The main lesson of this table is that the difference between the criteria can be small. In particular, it is negligible between total (or critical-level) utilitarianism and average utilitarianism. More surprisingly, it is also very small between the EDE and utilitarianism, but this is a direct consequence of adopting a utility function that varies very little between the consumptions of one and five units, so that the social priority between the two dynasties is mostly determined by the marginal utility, as in utilitarianism. It is an interesting observation that when describing the evolution of humanity over two million years, cross-section inequalities, and even inequalities between generations 1000 years apart, look small compared to longitudinal inequalities over the very long run. One could of course study other utility functions that have a different shape, but we chose here to retain the most popular functional form. Note also that the risk introduced in our bad scenario is quite benign as far as consumption is concerned, given the low growth rate in the favorable scenario.

The discount rates tend to zero in the table for the largest horizons, because the difference in consumption between the donors (5 units) and the worst-off in the worst scenario (1 unit) becomes very small in terms of annual growth rate. Even in the medium run, the discount rates are far from dismal. Note, however, that the risk introduced in our bad scenario is quite benign as far as consumption is concerned, given the low growth rate in the favorable scenario. Moreover, as we take a critical level equal to the EDE, the size of the population is not a very important parameter and therefore the loss of about half of the population is not considered a large loss for this criterion (but other options are possible which are more populationist). We do not want to defend these specific value judgments; instead we only aim at showing what ethical issues need to be addressed for a rigorous evaluation of climate policies.

Horizon	EDE	Total or CLU	Average U
50	-5,0140%	-4,9990%	-4,9993%
100	-2,5386%	-2,5308%	-2,5311%
200	-1,2768%	-1,2727%	-1,2730%
500	-0,5119%	-0,5101%	-0,5104%
1000	-0,2556%	-0,2546%	-0,2549%

Table 5: Discount rate for the EDE criterion, total or critical-level utilitarianism, and average utilitarianism

We now consider a variant, in which the 20% chance of the bad scenario is reduced to 19.9% and there is a 0.1% risk of a serious catastrophe that brings consumption back to a fraction of the current consumption level for all the future generations until the end after one million years of human history. (We do not change the population size in order to avoid an extra complication in the interpretation of the results.)

Table 6 presents the discount rate for the 50-year horizon depending on the extent of the catastrophe, as measured by the fraction of current consumption that is consumed in the future. Here again, the EDE criterion has a lower discount rate than the two utilitarian criteria because of its extra aversion to inequalities. The average utilitarian criterion has a lower discount rate than the total (or critical-level) utilitarian criterion because the marginal utility in the bad scenarios belongs to a larger fraction of the total population – the population is smaller than in the good scenario.

We now also present the discount rate for the additive criterion $\frac{1}{n}\sum_i \varphi\big(u(c_i)\big)$, because it is noteworthy that this criterion exhibits the "dismal" behavior discussed earlier in this paper. When the level of consumption tends to $\hat{c}$, marginal utility tends to the finite value $\hat{c}^{-2}$, but the marginal social value of $\varphi\big(u(c_i)\big)$ tends to infinity because $u(c_i)$ tends to zero, which generates very low discount rates. The correlation term that EDE adds to the discount rate eliminates this problem, by producing values of the discount rate that are closer to the marginal utility of the worst off. As explained earlier, this reflects the property of the EDE criterion that it respects risk preferences of the population when inequalities are limited.

Fraction	EDE	Total or CLU	Average U	Additive
1/10	-5,32%	-5,18%	-5,30%	-5,34%
1/50	-8,44%	-7,35%	-8,12%	-8,99%
1/100	-11,18%	-9,35%	-10,36%	-12,69%

1/190	-14,67%	-11,62%	-12,56%	-22,35%
1/199	-15,01%	-11,78%	-12,72%	-29,31%
1/199,999	-15,05%	-11,80%	-12,74%	-46,39%

Table 6: Discount rates when the additional catastrophic scenario has a fraction of current consumption.

These results nevertheless show that the small probability (1/1000) of a catastrophic scenario may significantly change all the discount rates even when they do not exhibit the dismal behavior and converge to a value that is well above -100%.

Conclusion

Let us briefly wrap up the argument of this paper. The discount rate only serves to measure the relative social priority of different individuals belonging to different generations. Therefore there is no need to worry about comparing the discount rate to the market interest rate, as a rational evaluation in terms of present value at the chosen discount rate will never fail to avoid dominated investments, and never fail to choose those with the greatest rate of return.

The key message of this paper is that discount rates are really to be computed between individuals (person-to-person), which gives a great role to inequalities within and between generations. In the very long run, Weitzman's observation that the worst-case scenario drives the discount rate has to be supplemented by the fact the situation of the worst off at both ends of the investment will also drive the value of the discount rate.

Therefore, if climate policies such as mitigation efforts are paid by the affluent populations of the present generations and greatly benefit the worst off of the distant future generations in the most catastrophic scenarios, it is very likely that the correct discount rates for the evaluation of such policies should be negative, which means that a dollar of benefit in the distant future is worth more than a dollar of effort today.

In conclusion, we would like to recall that a rigorous evaluation of climate policies is particularly challenging because it requires rethinking the welfare economics of risk, time, and population. In such endeavor the utilitarian criterion, which remains prominent in the debates about discounting, should be questioned and, perhaps, replaced with other criteria that better combine a certain respect for the risk preferences of the population and a substantial degree of aversion to inequality.

Finally, we should recall a point that has been made already by Stern (2007). The discount rate is useful to evaluate small transfers of consumptions across individuals living at different times. It is not the all-purpose tool that can serve for all evaluations. It is not adapted to large scale changes, and it is also not adapted to evaluating policies that change the size of the population or the probabilities of different scenarios. For such policies one has to go back to the underlying social welfare criteria. This is an additional reason to pay attention to the selection of such criteria on sound ethical principles.

To illustrate, consider that in the simulations of the previous section, the 4000th generation (the present) taxes the rich dynasty by 4 units (bringing its consumption, for one generation, to the level of other generation). For what reduction in the probability of the bad scenario (which is at 20% in the simulations) is this sacrifice worth making? For total utilitarianism, there is a big difference in social welfare between the two scenarios, so that a probability shift of 50 billionth would justify it. With a critical level above zero, this magnitude would rise. For average utilitarianism as well as the EDE criterion (with EDE critical level), for which the population loss in the bad scenario is not important, the probability shift should be at least 13 millionth, a much greater figure. This example shows that depending on the policy question, the similarity between criteria may vary substantially.

References

Arrow K.J., W.R. Cline, K.-G. Mäler, M. Munasinghe, R. Squitieri, J.E. Stiglitz 1996, "Intertemporal equity, discounting, and economic efficiency," in J.P. Bruce, H. Lee, E.F. Haites (eds.), *Climate Change 1995. Economic and Social Dimensions of Climate Change. Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press.

Fleurbaey M. 2010, "Assessing risky social situations", *Journal of Political Economy* 118: 649-680.

Fleurbaey M., S. Zuber 2012, "Intergenerational equity, risk, and discounting," mimeo.

Gollier C. 2002, "Discounting an uncertain future", *Journal of Public Economics* 85: 149-166.

Nordhaus W.D. 2007, "The Stern Review on the Economics of Climate Change," *Journal of Economic Literature* 45: 686–702.

Posner E.A., D. Weisbach 2010, *Climate Change Justice*, Princeton : Princeton University Press.

Schelling T. 1995, "Intergenerational discounting", *Energy Policy* 23: 395-401.

Stern N. *et al.* 2007, *The Economics of Climate Change*, New York: Cambridge University Press.

Weitzman M. 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. 2009, “On modeling and interpreting the economics of catastrophic climate change”, *Review of Economics and Statistics* 91: 1-19.