Discounting the future: the case of climate change.

Ivar Ekeland
Canada Research Chair in Mathematical Economics, UBC

September 17, 2009

Abstract

Acknowledgement 1 *These notes originate from a series of five lecture I gave at the PIMS Summer School on "Perceiving, measuring and managing risk" held at UBC, June 30 - July 11,2008, and from a Bachelier course held in Paris in 2008/09. The support of PIMS and of the Chaire de Développement Durable is gratefully acknowledged.. Copyright remains with the author, and these notes are not to be quoted without the authors permission*

1 Long-term policy-making.

1.1 What is the long term ?

These lectures deal with the economic aspects of long-term policy-making. As the historical notes will show, this problem has been around for many years, in fact since the beginnings of economic theory, and there is a vast literature on this subject. To state it as simply as possible, a decision-maker, either an individual or a collective entity (the government), is to make a decision today (time $t = 0$), the consequences of which will kick in only at a (much) later time $t = T$, and he/she has to weigh the immediate benefit of that decision against the future costs. Alternatively, the cost occurs today and the benefits at time T, and the question is then how much cost the decision-maker is willing to bear today in order to reap the benefits at time T.

What has changed, however, since these early days, is the time horizon. Up to very recently, what the economist, the engineer or the politician would consider long-term would be in the range 10 to 30 years (note for instance that the US Treasury does not issue bonds with a maturity longer than 30 years). Anything beyond that was considered beyond the horizon - just like accelerating galaxies slip beyond the boundary of the observable universe. This has changed in recent years, where the consequences of our actions beyond that horizon have become part of the agenda. Here are two examples:

- the lifetime of a nuclear plant is 40 to 60 years, after which it will have to be decommissioned and the site reclaimed, at a considerable cost, which has to be factored in the investment decision
- the Stern Review on Climate Change states that the course for the next 50 years is set: the inertia of the physical and biological system governing the Earth climate is such that the consequences of any policy we enact today will not be felt before 50 years have elapsed. The question is what happens after that, and the Stern Review depicts alternate scenarios spanning the 50 to 200 years period.

In these lecture notes, we will define the long term as the 50 to 200 years range.

1.2 What are the difficulties of long-term policy-making ?

There are two main features which set long-term decisions apart from short and middle-term ones. The first one is *high uncertainty.* This comes in two different guises:

- the predictable outcomes, that is those for which probabilities can be set, have a very high dispersion. For instance, the Stern Review states that, with a probability of 95%, under the business as usual scenario, the loss of GDP to the world econony 200 years from now will be in the range 2% to 35%
- but there also non-predictable events, which Stern calls "bifurcations", such as the cessation or the thermo-haline convector which runs the Gulf Stream. or the melting of the West Antarctic ice sheet. We are in no position to assign probabilities to such event but we know (a) that they can occur,and (b) that their consequences would be catastrophic

Let me mention "en passant" that the fact of global warming as such no longer is part of the uncertainty: it is now certain that it is occuring. At the time of this writing, there is a 50% chance that, for the first time in recorded history, the North Pole will be on open water.

The second difficulty is *non-commitment.* Whatever policy we enact today will presumably have to be adhered to until the desired consequences are achieved, 50 to 200 years from now, when we (or whoever has decided on these policies) no longer is there to carry them out. This means that we have to rely on future generations (and future governments) to carry them out when we are gone. There is no way we can commit unborn generations and whoever rules the planet one hundred from now to anything: they will do as they please. Whatever policy we design now for the long term has to answer the question: one hundred years from now, when the powers that be are supposed to implement these policies, is there a reasonable chance that they will do it ?

1.3 Is economic theory relevant ?

There are, of course, the economic approach to these problems is not the only possible or relevant one. Clearly, there are ethical considerations. As Keynes famously said, in the long term we are all dead. Do we care what happens after that ? Some people don't: this is the "après moi le déluge" philosophy, which has quite a number of proponents in academic and government circles. Most people do because of ethical considerations towards future generations and the planet itself. Adam Smith himself wrote that "the Earth and the fullness thereof belongs to all generations". There are also political agendas, with immediate gains or losses for decision-makers which preempt any long-term concern.

Even in the presence of ethical or political considerations, traditional cost-benefit analysis has shown itself to be a useful tool, if only to clarify the issues. To attach figures to policies does make a lot of difference. Part of the impact of the Stern Review is due to the fact that it came up with the conclusion that, under a policy of business as usual, climate change would cost 10% of GDP per year, while the cost of prevention stood at less than 1%. This is the kind of argument to which politicians and business leaders pay attention, and the scientific community should make every effort to speak to them in their own language.

Even so, applying cost-benefit analysis to long-term policies is by no means straightforward. In practice, this means discounting future benefits at a certain rate, say r. A natural choice for r would be the market rate of interest, especially for the longest maturity, which today stands at 4.6% (rate of 30-years US Treasury bond at the time of writing). In the following table, we give the present value of $1,000,000$ $ at $50,\ 100$, and 200 years, for interest rates of 10%, 4.6% and 1.4%, which is the value that the Stern Review took:

	50 ys	100 ys	200 ys
10%	$8,519$	73	< 0.00
4.6%	$105,540$	$11,140$	124
$1,4\%$	$499,000$	$249,000$	$62,000$

Clearly, an interest rate of 10% just wipes out the long term. The current market rate of 4.6% does somewhat better, but the rate of $1,4\%$ really make future events loom large. It is evident that the conclusions of the Stern report heavily depend on this choice of the interest rate, and that they would have been entirely different if, for instance, Stern had chose to discount at market rates. So the question now is: what justification, if any, is there for choosing such a low rate ? This is the question which we will address now.

1.4 Some historical notes

From the beginning of economics as a separate science, it was apparent that individual choices between present and future rewards were driven by some kind of time preference: individuals prefer to enjoy goods sooner rather than later. That theme was developed by John Rae (1834, [29]), Boehm-Bawerk (I884 [**?**]), Irving Fisher (1930, [14]), as a psychological trait of human nature. In 1960, Tjalling Koopmans [22] showed that impatience can be derived from benign assumptions on preferences. In other words, preferences in economic theory are usually ascribed to immediate consumption c, resulting in utility functions $u(c)$. If one now considers consumption schedules, $c(t)$ for $t \geq 0$ (possibly discrete), and tries to write down a reasonable set of axioms that preferences should satisfy, one is inevitably led to time preference as a logical consequence.

Since time preference is well established, the next question is how to translate it into a mathematical model. It is a natural idea to discount future utilities at a constant rate $\rho > 0$: the larger ρ, the more impatient the consumer. If $\rho = 0$, the consumer is indifferent between immediate and deferred consumption; in other words, he exhibits no impatience at all. The idea of setting up the question of economic growth as an optimisation problem is due to Frank Ramsey (1928, [30]). Interestingly, he chose $\rho = 0$ as his preferred option: "we do not discount later enjoyments in comparison with earlier ones, a practice that is ethically indefensible and arises merely from the weakness of the imagination". He did, however, treat the case $\varrho > 0$ as well, and that became the standard of the industry, following Samuelson (1937, [31])

1.5 Structure of the paper

We will proceed by first giving a detailed exposition of the standard model of economic growth. This is the topic of the next section; this model has very stringent assumptions (a single good in the economy, and constant discount rate), but as a result explicit formulas can be derived, including a formula for the long-term interest rate:

$$r = \rho + \theta g$$

which serves as the basis for much of the present discussions among economists. This formula is derived and explained in Section 3, and then discussed; suitable modifications are proposed to take into account the known deficiencies of the standard model and some of the current discussions about environmental issues and climate change. Many of these modifications lead us to non-constant interest rates.

Finally, we turn to concerns of intergenerational equity. In the standard model, there is a single infinitely-live individual, who stands as a proxy for people alive today and all future generations, which is clearly very simplistic. Following an idea of Sumaila and Walters [32], we separate the utility of people alive today and that of future generations, and we aggregate them in a single intertemporal criterion. This leads to a model similar to the standard one, but with non-constant interest rate.

At the end of this investigation, we have built an overwhelming case for considering non-constant interest rates, and this is the topic of the last section. It turns out that handling time-varying interest rates requires a change of paradigm and a new mathematical theory. Indeed, non-constant interest rates lead to time inconsistency, so that the notion of optimality changes with time.

The problem then becomes a game between successive generations, which was solved by Ekeland and Lazrak [10], [11], and we describe the main results of these papers.

2 The standard model of economic growth

2.1 Firms, consumers and growth

This model originates with the seminal paper of Ramsey in 1928. It is described in the opening chapters of most graduate textbooks in macroeconomics: see for instance [6] and [34]. or [4] for a more detailed treatment. There is *a single good* in the economy, which can be either consumed (in which case it is denoted by C) or used to produce more of the same good (in which case it is called capital and denoted by K) . This good is produced by a large (fixed) number of identical firms in perfect competition, so that they function as a single firm which is a price-taker. The global production function is:

$$Y = F(K, AL)$$

where K is the total capital invested in the economy, L is the labour force, and A is the productivity of labour (which will eventually depend on time, $A(t)$, to reflect technological progress). It will be assumed that there are *constant returns to scale*, so that setting $y = Y/AL$ (production per unit of effective labour) and $k = K/AL$ (capital per unit of productive labour), we have:

$$y = f(k)$$

where $f(k) := F(k, 1)$ is the reduced production function. It will be assumed to be concave, increasing, and satisfying the Inada conditions:

$$\begin{aligned} k &\geq 0, \quad f(k) > 0, \quad f(0) = 0 \\ f'(0) &= +\infty, \quad f'(k) \longrightarrow 0 \text{ when } k \longrightarrow \infty \end{aligned}$$

The population consists of identical individuals. Total consumption is C; up to a constant factor, it is also the consumption of each individual. Using the same scaling for production and for consumption, we find that the relevant variable is $c = C/AL$, the consumption per unit of effective labour. The consumption per individual is $C/L = Ac$, and this is the variable which enters the individual's utility function u. In the sequel, we will take the following specification:

$$u(x) = \frac{x^{1-\theta}}{1-\theta} \text{ for } \theta > 0,\ \theta \neq 1 \tag{1}$$

$$u(x) = \ln x \text{ for } \theta = 1 \tag{2}$$

so that the utility of each individual alive at time t is:

$$u(A(t)\,c(t)) = \frac{A(t)^{1-\theta}}{1-\theta} c(t)^{1-\theta}$$

We will consider (and compare) different consumption scenarios $C(t)$. To do that, we will treat the population as a single [1], infinite-lived individual (the so-called *representative consumer*), who consumes $A(t)\,c(t) = C(t)/L(t)$ at time t. Note that this is not the total consumption of society at time t, but the average consumption of its members; the difference is quite significant, since we will assume exponential growth of the population.

This single, infinite-lived individual, has a *pure rate of time preference* ρ. This means that, given the choice between consuming $c(0)$ at time $t = 0$ and $c(t)$ at time $t > 0$, he/she will be indifferent if and only if:

$$u(c(0)) = e^{-\rho t} u(c(t))$$

Let us note right now (and we will expand on this later) that ρ is NOT an interest rate.

The economy is driven exogeneously by technological progress and population growth, both of which happen at constant rates g and n. We set:

$$\begin{aligned} A(t) &= A(0)\,e^{gt} \\ L(t) &= L(0)\,e^{nt} \end{aligned}$$

so that the utility of the representative consumer at time t is:

$$\frac{A(0)^{1-\theta}}{1-\theta} c^{1-\theta} e^{g(1-\theta)t}$$

2.2 The planner's problem

2.2.1 Statement

We now imagine a benevolent and omniscient planner, who wants to maximize the intertemporal welfare of the representative consumer. He/she will consider the following problem:

$$\begin{aligned} &\max_c \frac{1}{1-\theta} \int_0^\infty c^{1-\theta} e^{g(1-\theta)t} e^{-\rho t} dt \qquad \text{(Ramsey)} \\ &\frac{dk}{dt} = f(k) - c - (n+g)\,k, \quad k(0) = k_0 \\ &k(t) \geq 0, \quad c(t) \geq 0 \end{aligned}$$

[1] A slight variant of the model, which leads to the same results, consists of assuming that the population consists of N identical dynasties, each of which will be treated as a single, infinite-lived individual.

The second equation represent the balance equation between savings and consumption. It states that at every moment t, (scaled) production $f(k)$ is fully allocated between immediate (scaled) consumption c and (scaled) capital investment dk/dt, the correction term $(n+g)k$ being there to take into account growth of population and technological progress. Of course k_0 is the initial capital.

Solving this problem leads to the following result, which will be proved in the next subsections. Assuming that $\rho > g(1-\theta)$ (the pure rate of time preference is high enough) we find that:

(a) There is a single k_∞, called the equilibrium value of capital, which solves the following equation, known as the *golden rule*:

$$f'(k_\infty) = \rho + \theta g + n \tag{3}$$

(b) The problem has a single solution $k(t)$, which has the property:

$$k(t) \longrightarrow k_\infty \text{ when } t \longrightarrow \infty$$

(c) The corresponding consumption $c(t)$ along the optimal path also converges:

$$c(t) \longrightarrow c_\infty = f(k_\infty) - (n+g)k_\infty$$

In equilibrium, when $k(0) = k_\infty$ and $c(0) = c_\infty$, we have $C(t) = C(0)e^{(n+g)t}$, $K(t) = K(0)e^{n+g}$ and $C(t)/L(t) = A(0)e^{gt}$, so that total consumption and total production are growing at the rate $n+g$, while consumption per head is growing at the rate g.

2.2.2 Existence and uniqueness

The Ramsey problem appears as a control problem, where the control is $c(t)$ and the state is $k(t)$. Substituting $c = f(k) - (n+g)k - \frac{dk}{dt}$ in the integrand, we reduce it to a problem in the calculus of variations, where the only unknown is k:

$$\max_{k\in\mathcal{A}} \int_0^\infty \frac{1}{1-\theta}\left(f(k) - (n+g)k - \frac{dk}{dt}\right)^{1-\theta} e^{-\beta t}dt \tag{4}$$

$$k(0) = k_0, \quad k(t) \geq 0 \tag{5}$$

$$f(k) - (n+g)k - \frac{dk}{dt} \geq 0 \tag{6}$$

We have set $\beta := \rho - g(1-\theta) > 0$, and $\mathcal{A}$ is the set of admissible functions, will be discussed later. For the sake of convenience, we also set:

$$I(k) = \int_0^\infty \frac{1}{1-\theta}\left(f(k) - (n+g)k - \frac{dk}{dt}\right)^{1-\theta} e^{-\beta t}dt$$

Note that, since u and f are strictly concave, strictly increasing, and positive, I is a strictly concave function of k with values in $R \cup \{+\infty\}$. As a consequence, if the maximum is attained on $\mathcal{A}$, and $\mathcal{A}$ is convex, it is attained at a single point (which will usually depend on the set $\mathcal{A}$).

Does the optimal solution exist ? Of course, the answer will depend on the definition of admissibility, that is, on the choice of $\mathcal{A}$. An appropriate choice for $\mathcal{A}$ would incorporate both regularity conditions (how smooth is $k(t)$) and growth conditions (how does $k(t)$ behave when $t \longrightarrow \infty$). One would then consider a minimising sequence $k_i(t)$, show that a subsequence converges to some $\bar{k}(t)$, and prove that $\bar{k}(t)$ is the minimizer. This is called the direct method in the calculus of variations. Unfortunately, I do not know of this method being applied successfully in the case at hand. The main difficulty is that the integrand $u(c)$ has slow growth when $c \longrightarrow \infty$ (it grows like a power α, with $a < 1$); as a consequence, there is nothing to prevent $k(t)$ from being discontinuous at some point t_0, as long as the jump is downwards (in other words, the consumer can consume instantaneously a non-zero quantity of good.

In the sequel, we shall prove existence in the class $\mathcal{A} = \mathcal{C}^2$ by another route (the royal road of Caratheodory)

2.2.3 A necessary condition: the Euler-Lagrange equation

Suppose there is a solution $k(t)$ for $\mathcal{A} = \mathcal{C}^2$. The classical Euler-Lagrange equation then holds for $k(t)$ (see any texbook on the calculus of variations), yielding a second-order equation:

$$\frac{d}{dt}\left(-u'\left(f(k) - (n+g)k - \frac{dk}{dt}\right)e^{-\beta t}\right) = u\left(f(k) - (n+g)k - \frac{dk}{dt}\right)\left(f'(k) - (n+g)\right)e^{-\beta t}$$

(with $u(c) = (1-\theta)^{-1} c^{1-\theta}$), which is valid as long as we do not hit the boundaries, that is, for $k(t) > 0$ and $c(t) > 0$. We simplify this expression by setting $c := f(k) - (n+g)k - \frac{dk}{dt}$, so that it becomes:

$$\frac{dc}{dt}\frac{u''(c)}{u'(c)} + f'(k) - \rho - \theta g - n = 0 \tag{7}$$

We can reduce this second-order equation to a system of first-order equations:

$$\frac{dk}{dt} = f(k) - (n+g)k - c \tag{8}$$

$$\frac{dc}{dt} = -\frac{1}{\theta}\left(\rho + \theta g + n - f'(k)\right)c \tag{9}$$

All the solutions of this system can be represented on a two-dimensional *phase diagram*. The vertical line $\rho + \theta g + n = f'(k)$ and the curve $f(k) - (n+g)k = c$ divide the positive orthant R_+^2 in four regions, which have a common boundary point (k_∞, c_∞), where:

$$\begin{aligned} f'(k_\infty) &= \rho + \theta g + n \\ c_\infty &= f(k_\infty) - (n+g)k_\infty \end{aligned}$$

We can find out the sign of dk/dt and dc/dt in each region, and draw the phase diagram. It follows that (k_∞, c_∞) is the unique fixed point, and that it is unstable. There is a unique solution $\left(\bar{k}(t), \bar{c}(t)\right)$, which converges to that fixed point:

$$\bar{k}(t) \longrightarrow k_\infty \text{ and } \bar{c}(t) \longrightarrow c_\infty \text{ when } t \longrightarrow \infty$$

all the others go one of the boundaries, $c = 0$ or $k = 0$.

Of course, we strongly suspect that $\left(\bar{k}(t), \bar{c}(t)\right)$ is the optimal solution to the problem when $\mathcal{A} = \mathcal{C}^2$. We can try to compare it directly with the other solutions of the Euler-Lagrange equation (7), but this is quite ponderous (one has to investigate separately all the four regions in the phase space) and will not generalize to more general situations (for instance, with several goods). In the literature, one points points out that the Euler equation should be supplemented by two boundary conditions, one of which is known, namely $k(t_0) = k_0$, while the other should describe the behaviour of $k(t)$ when $t \longrightarrow \infty$. This is the celebrated "transversality condition at infinity", several versions of which have been given, either in discrete (see [[**?**]], [[12]]) or in continuous (see [[2]], [[24]], [**??**]) time. Unfortunately, none of them applies to the present situation, where the candidate solution is the only one which does not go to the boundary.

So we will prove that $\left(\bar{k}(t), \bar{c}(t)\right)$ is the optimal solution by a very different method, using the Hamilton-Jacobi equation, which we now introduce.

2.2.4 Another necessary condition: the Hamilton-Jacobi-Bellman equation

If there is an optimal solution for every initial point k_0, depending smoothly on k_0, the function:

$$V(k_0) := \max\left\{I(k) \mid k \in C^2,\ k(0) = k_0\right\}$$

satisfies the following relation at every point where it is differentiable:

$$V'(k) \geq 0 \text{ and } \frac{\theta}{1-\theta} V'(k)^{1-1/\theta} + V'(k)\left(f(k) - (n+g)k\right) - \rho V(k) = 0 \tag{HJB}$$

and the optimal consumption is given by a feedback strategy:

$$c(t) = V'(k(t))^{-1/\theta} \tag{10}$$

An informal proof is as follows. Define:

$$V(k_0) := \max_{c(\cdot)}\left\{\int_0^\infty u(c(t))\, e^{-\rho t} dt \mid k(0) = k_0\right\}$$

(with $u(c) = (1-\theta)^{-1} c^{1-\theta}$). Using optimality, we have:

$$\begin{aligned}
V(k_0) &= \max_{x,c(\cdot)} \left\{ \varepsilon u(x) + \int_{\varepsilon}^{\infty} u(c(t)) e^{-\rho t} dt \mid k(\varepsilon) = k_0 + \varepsilon (f(k_0) - x - (n+g) k_0) \right\} \\
&= \max_{x,c(\cdot)} \left\{ \varepsilon u(x) + e^{-\rho\varepsilon} \int_{0}^{\infty} u(c(t)) dt \mid k(\varepsilon) = k_0 + \varepsilon (f(k_0) - x - (n+g) k_0) \right\} \\
&= \max_{x} \left\{ \varepsilon u(x) + \ e^{-\rho\varepsilon} V(k_0 + \varepsilon (f(k_0) - x - (n+g) k_0)) \right\} \\
&= \max_{x} \left\{ \varepsilon u(x) + \ (1-\rho\varepsilon)(V(k_0) + V'(k_0)(f(k_0) - x - (n+g) k_0)) \right\} \\
&= V(k_0) + \varepsilon \max_{x} \left\{ u(x) - xV'(k_0) - \rho V(x_0) + V'(k_0)(f(k_0) - (n+g) k_0) \right\}
\end{aligned}$$

We end with:

$$0 = \max_{x} \left\{ u(x) - xV'(k_0) - \rho V(x_0) + V'(k_0)(f(k_0) - (n+g) k_0) \right\} \tag{11}$$

The right-hand side splits in two terms, yielding the desired equation:

$$0 = \tilde{u}(V'(k_0)) - \rho V(k_0) + V'(k_0)(f(k_0) - (n+g) k_0)$$

where:

$$\begin{aligned}
\tilde{u}(y) \quad &: = \max_{x} \{ u(x) - xy \} \\
&= \max_{x} \left\{ \frac{1}{1-\theta} x^{1-\theta} - xy \mid y = x^{-\theta} \right\} \\
&= \left(\frac{1}{1-\theta} - 1 \right) x^{1-\theta} = \frac{\theta}{1-\theta} y^{-\frac{1-\theta}{\theta}}
\end{aligned}$$

and the maximum is achieved for $y = u'(x) = x^{-\theta}$, yielding $x = y^{-1/\theta}$, and hence formula (10).

2.2.5 A sufficient condition: the royal road of Caratheodory

Theorem 2 *Suppose (HJB) has a C^2 solution $V(k)$ such that, for any $k_0 > 0$, the solution $\bar{k}(t)$ of the Cauchy problem:*

$$\frac{dk}{dt} = f(k) - (n+g) k - i(V'(k)), \ k(0) = k_0 \tag{12}$$

converges to k_∞ and $e^{-\beta t} V(\bar{k}(t)) \longrightarrow 0$ when $t \longrightarrow \infty$. Then, for any starting point $k_0 > 0$, the path given by (12) is optimal among all C^2 paths $k(t)$ such that

$$\lim \sup_{T \longrightarrow \infty} e^{-\beta t} V((k(T))) \geq 0$$

In particular, it is optimal among all interior paths (that is, paths along which $k(t)$ is bounded away from 0 and ∞).

Proof. Consider any path $c(t), k(t)$ starting from k_0. Because of equation (11), we have:

$$\int_0^T e^{-\beta t}\left[u(c(t)) - c(t)V'(k(t)) + V'(k(t))(f(k(t)) - (n+g)k(t) - \beta V(k(t)))\right]dt \leq 0 \text{ for every } T > 0$$

The left-hand side can be rewritten as follows:

$$\begin{aligned}
\int_0^T e^{-\beta t}u(c(t))\,dt + \int_0^T e^{-\beta t}\left[\frac{dk}{dt}V'(k(t)) - \beta V(k(t))\right]dt &\leq 0 \\
\int_0^T e^{-\beta t}u(c(t))\,dt + e^{-\beta t}V(k(t))\,|_0^T &\leq 0 \\
\int_0^T e^{-\beta t}u(c(t))\,dt + \lim\sup_{T\longrightarrow\infty} e^{-\beta t}V((k(T))) &\leq V(k(0))
\end{aligned}$$

Letting $T \longrightarrow \infty$, we find:

$$\int_0^\infty e^{-\beta t}u(c(t))\,dt \leq V(k_0)$$

. On the other hand, setting $\bar{c}(t) = i\left(V'\left(\bar{k}(t)\right)\right)$, we get the path $\bar{k}(t)$ and the equality is achieved because of equation (HJB). Hence the result. ■

Note that the C^2 regularity of $V(k)$ everywhere (including at $\bar{k}$) has played a crucial role in the proof

Theorem 3 *There is a C^2 solution of the HJB equation such that all solutions of (12) converge to the point k_∞ defined by:*

$$f'(k_\infty) = \rho + \theta g + n$$

This will be proved much later in the lectures. As a consequence, we have:

Corollary 4 *The Ramsey problem has a unique solution $k(t)$ in the class C^2, with $k(t) \longrightarrow k_\infty$ when $t \longrightarrow \infty$.*

2.2.6 Local analysis

Linearizing equations (8) (9) near the stationary point, with $y = k - k_\infty$ and $x = c - c_\infty$, we get:

$$\begin{aligned}
\frac{dy}{dt} &= (f'(k_\infty) - (n+g))y - x = \rho y - x \\
\frac{dx}{dt} &= f''(k_\infty)\frac{c_\infty}{\theta}y
\end{aligned}$$

where we have taken into account the fact that $\frac{dc}{dt} = 0$ at the stationary point. The corresponding matrix is:

$$\begin{pmatrix} \rho & -1 \\ \frac{c_\infty}{\theta}f''(k_\infty) & 0 \end{pmatrix}$$

which has two real roots, one positive and one negative. Along the "optimal" trajectory, we have:

$$\left(\begin{array}{c} k-k_{\infty} \\ c-c_{\infty} \end{array}\right) \sim \exp\left[\frac{1}{2}\left(\rho-\sqrt{\rho^{2}+\frac{4}{\theta}c_{\infty}f''\left(k_{\infty}\right)}\right)t\right]\left(\begin{array}{c} k\left(0\right)-k_{\infty} \\ c\left(0\right)-c_{\infty} \end{array}\right)$$

2.3 The equilibrium problem

What if there is no planner, or if the planner has no means of implementing his/her policy ? In that case, we will be looking for an equilibrium interest rate, that is, an interest rate $r(t)$ for which markets clear. By definition, if the spot rate is $r(t)$, the price today of one unit of consumption available at time t is $\exp(-R(t))$, where:

$$R(t) = \int_0^t r(s)\, ds$$

Note that, if the interest rate is constant, $r(t) = r$, we find the usual formula, $R(t) = \exp(-rt)$, but, as we shall see, there is no particular reason that this should be the case in equilibrium.

Assume the yield curve is $r(t)$, which is common knowledge and let us write the market-clearing conditions. There is only one because the representative consumer and the representative firm enter forward contracts at time $t = 0$. On the supply side, the interest rate (spot rate on the money market) should be equal to the *marginal return on investment* (between t and $t + dt$):

$$r(t) = f'(k(t)) - n - g$$

Since the firm makes no profit, its revenue is shared between labor and capital, so that the wage must be:

$$w(t) = f(k) - kf'(k) + n + g$$

Recall that the representative citizen consumes $A(t)\,c(t) = A(0)\,e^{gt}c(t)$ at time t. He/she maximizes intertemporal utility:

$$\max_c \int_0^{\infty} \frac{1}{1-\theta} c(t)^{1-\theta}\, e^{-\beta t} dt$$

(with $\beta = \rho - (1-\theta)\,g$, as above) subject to the budget constraint (intertemporal borrowing and lending is allowed):

$$\int_0^{\infty} e^{-R(t)} c(t)\, e^{(g+n)t} dt \leq k(0) + \int_0^{\infty} e^{-R(t)} w(t)\, e^{(g+n)t} dt$$

Introducing a Lagrange mutiplier λ, this problem becomes (with $u(c) = (1-\theta)^{-1} c^{1-\theta}$):

$$\max_c \int_0^{\infty} \left[u(c(t))\, e^{-\beta t} + \lambda e^{-R(t)} \left[w(t) - c(t)\right] e^{(n+g)t}\right] dt \qquad (13)$$

where $R(t)$ and $w(t)$ are known. Note for future reference that the exponent $R(s)$ is non-constant. The optimal solution to the consumer's problem, as seen from time $t = 0$, is:

$$u'(c(t))\, e^{-\beta t} = \lambda e^{-R(t)+(n+g)t}$$

Let us transform this equation a little bit:

$$\begin{aligned} \ln u'(c(t)) - \beta t &= \ln \lambda - R(t) + (n+g)\, t \\ -\frac{\theta}{c}\frac{dc}{dt} - \beta &= n + g - f'(k) \end{aligned} \tag{14}$$

This is just the Euler-Lagrange equation, which is satisfied by the solution to the planner's problem. Hence:

Proposition 5 *The solution to the planner's problem is also a solution to the equilibrium problem. They are both efficient (Pareto optimal).*

There is another way to retrieve the Euler-Lagrange equation. Consider *the marginal return on consumption*. In equilibrium, it should be equal to the interest rate (and hence to the marginal return on investment). More precisely, consider the representative consumer at time $t = 0$, and his/her intertemporal consumption path $\tilde{c}(t) = e^{gt}c(t)$; let us ask ourselves how much consumption $\Delta\tilde{c}(0)$ he/she would be willing to forgo at time 0 in other to increase his/her consumption by $\Delta\tilde{c}(t)$ at time t. The balance equation can be written (with $u(c) = (1-\theta)^{-1} c^{1-\theta}$ and $\beta = \rho - (1-\theta)\, g$)

$$\begin{gathered} e^{-\rho t} u'(\tilde{c}(t))\, \Delta\tilde{c}(t) - u'(\tilde{c}(0))\, \Delta\tilde{c}(0) = 0 \\ \frac{\Delta\tilde{c}(t)}{\Delta\tilde{c}(0)} = \frac{u'(\tilde{c}(0))}{u'(\tilde{c}(t))} e^{\beta t} = \exp \int_0^t r(s)\, ds \end{gathered}$$

with

$$r(t) = \beta - \frac{d}{dt} \ln u'(\tilde{c}(t)) = \beta + \frac{\theta}{\tilde{c}}\frac{d\tilde{c}}{dt}(t)$$

Writing $r(t) = f'(k(t)) - (n+g)$, and plugging in $\tilde{c}(t) = e^{gt}c(t)$, so that:

$$\frac{1}{\tilde{c}}\frac{d\tilde{c}}{dt} = g + \frac{1}{c}\frac{dc}{dt}$$

we find the Euler-Lagrange equation. It expresses that, in equilibrium, the interest rate in the economy is equal to the marginal return on investement, and also to the marginal return on consumption.

2.4 Bibliography

As mentioned in the introduction, the standard model can be found in the textbooks, for instance [6] , [34] or [4], although not in a way that would fully satisfy a mathematician. The problem of finding the right transversality condition at infinity is still open: there are many versions around (see [2], [12], [24], [20]), but none which applies to the standard model (where the optimal solution is isolated, all the other ones being either unbounded or hitting the boundary). For the Hamilton-Jacobi-Bellman equation, and an introduction to the royal road of Caratheodory, see [8] or [7]

3 Determinants of the interest rate.

We shall use the standard model as a benchmark, and introduce successive modifications.

3.1 The classical theory

In the preceding section, we have proved that in the framework of the standard model, where:

- the rate of growth of average consumption is constant and equal to g,
- the utility function of the representative consumer is $u(c) = (1-\theta)^{-1} c^{1-\theta}$ with $\theta > 0$ (CRRA: constant relative risk aversion)

the equilibrium interest rate in the economy is given by:

$$r(t) = f'(k(t)) - n - g \tag{15}$$

$$= \rho - (1-\theta) g + \frac{1}{c(t)} \frac{dc}{dt} \tag{16}$$

Here $(k(t), c(t))$ is the optimal (from the point of view of the planner) or equilibrium (from the point of view of the representative consumer) scenario for the economy, starting from $k(0) = k_0$. The equality in (15) and (16) expresses that in equilibrium, the marginal return on consumption equals the marginal return on investment.

Note an important consequence of these formulas: *the spot rate $r(t)$ at time t is not constant* (even though all the other parameters in the economy, including the psychological discount rate, is constant). More precisely, consider the yield curve at time $t > 0$. This is the map $T \longrightarrow r_t(T)$, defined for $T > t$, with:

$$r_t(T) := \frac{1}{T} \ln \int_t^T r(s)\, ds$$

so that $\exp r_t(T)$ is the price at time t of one unit of numéraire delivered at time $T > t$. We find that:

- the yield curve is not flat
- it changes with time
- as $t \longrightarrow \infty$ it converges to a flat yield curve, $r_t(T) \longrightarrow r_\infty$

- for every t, the long-term rate at time t is $\lim_{T\longrightarrow\infty} r_t(T) = r_\infty$

Here r_∞ is the spot rate at the stationary point (k_∞, c_∞). At this point, formulas (15) and (16) become:

$$\begin{aligned} r_\infty &= f'(k_\infty) - n - g && (17)\\ &= \rho + \theta g && (18) \end{aligned}$$

Note the remarkable fact that it no longer depends on t. In other words, in the standard model, *the long-term interest rate is constant* along the optimal path and equal to the spot rate at the stationary point.

Formulas (16) and (18) generalize to a much broader class of models than the standard one, provided there is one consumption good and one representative consumer. Consider an infinite-lived individual, with utility function u, pure rate of time preference $\rho > 0$, and who is facing a schedule of consumption $c(t)$, leading to an overall utility of:

$$\int_0^\infty e^{-\rho t} u(c(t))\, dt$$

Let us ask ourselves how much immediate consumption $\Delta c(0)$ he/she would be willing to forgo in order to increase its consumption by $\Delta c(t)$ at some later time $t > 0$. Assuming these are small quantities, we can work on the margin, and we get the relation:

$$\frac{\Delta c(t)}{\Delta c(0)} = \frac{u'(c(0))}{u'(c(t))} e^{\rho t}$$

so that the marginal return on consumption is:

$$\frac{\Delta c(t)}{\Delta c(0)} = \exp \int_0^t r(s)\, ds$$

In equilibrium, if we use the consumption good as numéraire, this should be equal to the spot interest rate:

$$r(t) = \rho - \frac{d}{dt} \ln u'(c(t)) = \rho - \frac{u''(c(t))}{u'(c(t))} \frac{dc}{dt}(t)$$

This is most conveniently rewritten as follows:

$$\begin{aligned} r(t) &= \rho + \left(-c(t) \frac{u''(c(t))}{u'(c(t))} \right) \left(\frac{1}{c(t)} \frac{dc}{dt}(t) \right) \\ &= \rho + \eta(c(t))\, G(t) && (19) \end{aligned}$$

where:

- $\eta(c) := -cu''(c)/u'(c)$ is a positive parameter (because u is concave), usually called the relative risk aversion; in this context, it would be more relevant to call it the relative satiation. It usually depends on the level of consumption c. In the special case of power utilities, $u(c) = c^{1-\theta}/(1-\theta)$, it is constant and equal to θ

- $G(t) := (dc/dt)/c(t)$ is the rate of growth of the economy

Formula (19) is the benchmark for determining the interest rate, and is generally accepted in the economic literature. For instance, the Stern report takes $\rho = 0.1\%$, $\theta = 1$ and $g = 1.3\%$, yielding $r = 1.4\%$. Most of its critics claim that it is too low, and take $\rho = 2\%$, $\theta = 2$ and $g = 2\%$ as more reasonable numbers, yielding $r = 6\%$. We will discuss these claims, and bring more economic arguments to bear, in the sequel. Meanwhile, let us make some observations:

- as soon as there is growth in the economy ($G(t) > 0$), we have $r > \rho$. For instance, as Ramsey found out, we can have positive interest rate $r > 0$ even if the pure rate of time preference is zero, $\rho = 0$.

- the interest rate *rises* with the growth rate g. For instance, setting $\rho = 2\%$ and $\theta = 2$, we get $r = 6\%$ if $g = 2\%$ and $r = 10\%$ if $g = 4\%$. Why is that so ? Well, ask yourself the following question. Historically, growth has been around 2% for the past two hundred years. Now, imagine how your own ancestors were living 200 years ago - probably in conditions which you would consider of extreme need and poverty. Would you want such miserable people to have set something aside for you ? Probably not - quite the opposite, if you were able to do something for them, you would do it. Well, it growth continues at the same rate, this is the way that our descendants will look upon us; they will be richer than we can imagine. So why should we make sacrifices for such people ? Hence the high interest rate that we are in fact charging them.

- on the other tack, the interest rate *falls* with the growth rate. For instance, setting $\rho = 2\%$ and $\theta = 2$ again, we get $r = -2\%$ if $g = -2\%$, that is, if the economy contracts at the rate of 2% a year. So *negative interest rates* are not unthinkable - they might actually be needed in periods of negative growth. Think for instance of an economy where the only good is the environment, which cannot be produced, and actually has to decrease as the population growth - in such an economy, the interest rate would have to be negative. This leads us to the idea that one would actually have to use different rates for environmental goods and for consumption (manufactured) goods. The proper setting for exploring this idea is a two-goods model, and this is what we will be doing next.

3.2 Modifications 1: The environment as a separate good

This section develops, in a continuous-time framework, the ideas of Guesnerie [16], [17]We complement the standard model by adding a environment good E, along with the consumption good C. The two goods have different characteristics:

- E is a public good, and cannot be produced: it should be understood as the global quality of the environment.

- C is a private good, and can be produced as in the standard model: it should be understood as an aggregate of all consumption goods.

The consumption good will be used as numéraire. For the time being, we assume that the environment good is available in a fixed quantity $\bar{E}$ (so that the quantity will not be decreased as the economy grows).

Along the lines of Guesnerie, we choose the utility function of the representative consumer to be:

$$u(C,E) = \frac{1}{1-\theta} v(C,E)^{1-\theta}$$

with

$$v(C,E) = (C^{\alpha} + E^{\alpha})^{1/\alpha}$$

The parameter $\alpha \leq 1$ denotes the extent to which the environment good E and the consumption good C are substitutes. If a simultaneous and marginal changes $C \longrightarrow C - \Delta C$ and $E \longrightarrow E + \Delta E$ is to leave the total utility invariant, then we must have:

$$-\frac{\partial v}{\partial C}\Delta C + \frac{\partial v}{\partial E}\Delta E = 0$$

so that:

$$\frac{\Delta E}{\Delta C} = \frac{\partial v}{\partial C} / \frac{\partial v}{\partial E} = \left(\frac{E}{C}\right)^{1-\alpha}$$

The right-hand side can be rewritten as follows:

$$\frac{\Delta E}{E} = \left(\frac{E}{C}\right)^{-\alpha} \frac{\Delta C}{C}$$

In other words, to achieve an increase of 1% in the environmental good, the representative individual is willing to give up $\left(\frac{E}{C}\right)^{\alpha}$ % of the consumer good.

- if $0 < \alpha \leq 1$, the willingness to pay for the environmental good decreases as E/C decreases, that is, as it becomes relatively scarcer. This is the case when the environmental good and the consumption good are *substitutes*.

- if $\alpha < 0$, the willingness to pay for the environmental good increases as it becomes relatively scarcer. This is the case when the two goods are *complements*.

As in the standard model, we assume that there is a production function $Y = F(K, AL)$, which is positively homogeneous of degree one, and where the labour force $L(t) = L_0 \exp nt$ and the technological progress $A(t) = A_0 \exp gt$ are exogeneously given. Introducing the reduced consumption $c(t) = C(t) A(t)^{-1} L(t)^{-1}$, as in the preceding section, so that the average consumption at time t is $A(t) c(t)$, we find that the utility of the representative consumer at time t is given by:

$$u(c(t)) = \frac{1}{1-\theta} \bar{E}^{1-\theta} \left(1 + \left(\frac{A_0}{\bar{E}}\right)^{\alpha} c(t)^{\alpha} e^{g\alpha t}\right)^{(1-\theta)/\alpha}$$

The representative consumer's optimisation problem then becomes:

$$\max \int_0^\infty \frac{1}{1-\theta} \bar{E}^{1-\theta} \left(1 + \left(\frac{A_0}{\bar{E}}\right)^\alpha c(t)^\alpha e^{g\alpha t}\right)^{(1-\theta)/\alpha} e^{-\rho t} dt$$
$$\frac{dk}{dt} = f(k) - (n+g)k - c$$

3.2.1 The case $\alpha < 0$

In that case, we find that, for large t, the utility function can be approximated as follows:

$$\left(1 + \left(\frac{A_0}{\bar{E}}\right)^\alpha c(t)^\alpha e^{g\alpha t}\right)^{(1-\theta)/\alpha} \simeq 1 + \frac{1-\theta}{\alpha} \left(\frac{A_0}{\bar{E}}\right)^\alpha c(t)^\alpha e^{g\alpha t}$$

The constants play no role in intertemporal optimisation, and we are left with the criterion:

$$\max \int_0^\infty c(t)^a e^{(\rho-\alpha g)t} dt$$

This is precisely the standard problem again. We find that $c(t) \longrightarrow c_\infty$ and $k(t) \longrightarrow k_\infty$, where:

$$f'(k_\infty) = \rho + (1-\alpha) g + n, \quad c_\infty = f(k_\infty) - (n+g) k_\infty$$

In the absence of environmental concerns, the stationary level of capital would be k_∞^c, given by:

$$f'(k_\infty^c) = \rho + g + n < \rho + (1-\alpha) g + n$$

so that $k_\infty^c > k_\infty$. In other words, the presence of a non-substituable environmental good *lowers* economic growth. More precisely, the economy grows as $k_\infty \exp(g+n)t < k_\infty^c \exp(g+n)t$, and, of course, the amount of consumption is reduced accordingly. This is known in the literature as "ecological stunting".

The interest rate at the stationary point then is:

$$r_\infty = \rho + (1-\alpha) g > \rho + g$$

Note that the parameter θ does not appear in thes formulas - the risk aversion of the representative consumer does not come into play (at least not at the stationary state) ! The only relevant parameter is the substitution rate between the public and the private good. As technological progress and population growth drive up the production of consumer goods, the environment becomes comparatively more valuable, and long–term interest rates are determined only by ρ, the pure rate of time preference, g, the technological growth rate, and α - the larger α, the less an increase in consumption can compensate for a decrease in environment quality, and the higher the interest rate.

We can now ask ourselves whether an investment that will result in a one-time increase in consumption $\Delta c(0)$ today and result in a permanent decrease

ΔE in the quality of the environment is worth undertaking. The answer will be yes if and only if: that it is equal to:

$$u'(c(0))\,\Delta c(0) > \left[\frac{\partial}{\partial E}\int_0^\infty \frac{1}{1-\theta}\left(\bar{E}^\alpha + A_0^\alpha c(t)^\alpha e^{g\alpha t}\right)^{(1-\theta)/\alpha} e^{-\rho t}dt\right]\Delta E$$

$$\frac{c(0)^\alpha}{\left(c(0)^\alpha + \bar{E}^\alpha\right)^{1-1/\alpha}}\Delta c(0) > \bar{E}^{\alpha-1}\Delta E\int_0^\infty \left(\bar{E}^\alpha + A_0^\alpha c(t)^\alpha e^{g\alpha t}\right)^{(1-\theta-\alpha)/\alpha} e^{-\rho t}dt$$

On the stationary path, $c(t) = c_\infty$, we get:

$$\frac{\Delta c}{\Delta E} > \left(c_\infty^\alpha + \bar{E}^\alpha\right)^{1-1/\alpha}\frac{\bar{E}^{\alpha-1}}{c_\infty^\alpha}\int_0^\infty \left(\bar{E}^\alpha + A_0^\infty c_\infty^\alpha e^{g\alpha t}\right)^{(1-\theta-\alpha)/\alpha} e^{-\rho t}dt$$

3.2.2 The case $0 < \alpha \leq 1$

As $t \longrightarrow \infty$, we find that:

$$u(c(t)) \simeq \frac{1}{1-\theta}A_0^{(1-\theta)}c^{(1-\theta)}e^{(1-\theta)gt}$$

and we are back into the standard model.This time the environmental good simply disappears from the global picture.

3.3 Modifications 2: Uncertainty on the growth rates

3.3.1 A classical argument

Let us start from Ramsey's formula:

$$r = \rho + \theta g \tag{20}$$

Assume now that *we believe in the model*, but are uncertain about the growth rate g :

$$g \sim \mathcal{N}\left(\bar{g}, \sigma^2\right)$$

so that average consumption $c(t) = c(0)\,e^{gt}$ is lognormal.

Assume moreover that we are utility maximizers, and handle uncertainties à la von Neumann-Morgenstern. We ask, as always, how much consumption $\Delta c(0)$ we are willing to forgo today to increase by $\Delta c(t)$ our consumption at time t

$$u'(c(0))\,\Delta c(0) = E\left[u'(c(T))\,\Delta c(T)\right]e^{-\rho t}$$

with $u(c) = \frac{c^{1-\theta}}{1-\theta}$ so that $u'(c) = c^{-\theta}$. This gives:

$$\frac{\Delta c(0)}{\Delta c(t)} = \frac{e^{-\rho t}}{u'(c(0))}E\left[u'(c(t))\right]$$

The computation gives:

$$\frac{e^{-\rho t}}{c(0)^{-\theta}} E\left[c(0)^{-\theta} e^{-g\theta t}\right] = e^{-rt}$$

Using the well-known properties of the lognormal distribution, we find that:

$$E\left[e^{-g\theta t}\right] = \exp\left[-\bar{g}\theta t + \frac{1}{2}\theta^2\sigma^2 t^2\right]$$

and hence:

$$r = \rho + \theta\bar{g} - \frac{1}{2}\theta^2\sigma^2 t \tag{21}$$

Note that we are back with non-constant interest rates: in fact, the interest rate goes to $-\infty$ when t goes to ∞ ! It may be for this reason that the economic literature takes a different route, and simply averages the Ramsey formula (22), *defining* r as the mathematical expectation of the right-hand side. We then get the formula:

$$r = \rho + \theta\bar{g} - \frac{1}{2}\theta^2\sigma^2 \tag{22}$$

which, although simpler (we are back with constant interest rates) and popular (see [19] or makes little sense to me. If we take the usual values, $\rho = \theta = g = 2$, we get $r = 6\% - 2\sigma^2$, and since σ (the volatility of the growth rate) is of the order of a few percentage points, this correction will not be enough to reach the Stern value of 1.4%. On the other hand, formula (21) will drive the interest rate down to very low values, and eventually to negative ones: in the very long term, the uncertainty becomes so large that it forbids any risk-taking.

One thing is for sure: uncertainty *lowers* the interest rate. This corresponds to the standard fact that individuals are risk-averse. Note that this runs counter to an argument that politicians and companies have been making for many years, namely that we should do nothing about climate change, because it is not certain and it may turn out to be all right after all. From what we know, people are risk-averse for themselves, at least when the stakes (magnitude of potential losses) is large, meaning that the downside is more important to them than the upside, and it is difficult to understand why society should behave differently.

3.3.2 Pooling opinions of experts

In 2001, Weitzman [36] made the following, very general, observation. Suppose you consult two experts, whom you equally trust, about which interest rate to choose, and that they come up with two different opinions, namely r_1 and r_2, with $r_1 < r_2$. What value should you take ? As you trust them equally, it seems reasonable to pick the mean value, namely $\bar{r} = (r_1 + r_2)/2$. As Weitzman points out, this is wrong: what these experts are really saying is that one dollar today is worth respectively $e^{-r_1 t}$ and $e^{-r_2 t}$ at time $t > 0$. So if a mean is to be taken, it should be the mean of those values, leading to a interest rate $\tilde{r}$ given by:

$$\tilde{r}(t) = \frac{1}{t} \ln \left(\frac{1}{2} e^{-r_1 t} + \frac{1}{2} e^{-r_2 t} \right)$$

Note that this interest rate is *not constant*. It is approximately equal to $\bar{r} = (r_1 + r_2)/2$ for the short term, but for the long term is equal to the lowest rate r_1. This is the Weitzman lesson: for the long term, the lowest rate should prevail

Weitzman put his idea into practice. He pooled $I = 1,800$ economists and asked them for an assessment of interest rates to be applied for investment projects. Economist i answered with a constant rate r_i. leading to a discount rate $R_i(t) = e^{-r_i t}$. He found that the r_i were distributed according to a Gamma distribution with parameters (α, β):

$$f(r) \sim \frac{\beta^\alpha}{\Gamma(\alpha)} r^{\alpha - 1} e^{-\beta r}$$

Averaging the discount rates, he then derives the aggregate interest rate:

$$R(t) := \int_0^\infty f(r) A(r) \, dr = \left(\frac{\beta}{\beta + t} \right)^\alpha = \frac{1}{(1 + t\sigma^2/\mu)^{\mu^2/\sigma^2}}$$

In terms of the mean μ and the variance σ^2 of the Gamma distribution (α, β). The corresponding interest rate then is:

$$r(t) = -\frac{1}{R} \frac{dR}{dt} = \frac{\alpha}{\alpha + \beta} = \frac{\mu}{1 + t\sigma^2/\mu}$$

Note that very long-term interest rates are 0. The question is, how far out is the very long term ? Within the time horizon of the Stern review, from 50 to 200 years, Weitzman finds an interest rate of 1.75%, very much in line with the value 1.4% chosen by Stern himself.

3.4 Modifications 3: Uncertainty on the model.

Up to now, the modelling does not capture one of the main features of very-long term decisions, namely the possibility of major catastrophes with unknown probabilities. The fact that these probabilities are unknown is an added ingredient to risk, which is not captured by simply assigning a priori probabilities, as in classical economic theory. Indeed, a classicial experiment by Ellsberg [13] indicates that people have a specific aversion to ambiguity, that is to facing unknown probabilities. This is not captured by the von Neumann-Morgenstern approach to decision under uncertaintly, and the paradigm has to be changed. There is at present an active and promising literature on decision making under Ellsberg ambiguity (see [21], [15] for instance)

Weitzman [37] has pointed out another problem: whatever probability distribution our model works with, this will not be the one we work with. Indeed, we do not observe the distribution, all we can do is to infer it from a *finite* (and, in the case of climate change, pitifully small) amount of data. This means that, even if our model specifies Gaussian or Poisson distributions, which is usually the case, and which are nice because they have "thin tails" (large deviations have small probabilities) the ones we will end up working with may well have "fat tails", meaning that all long-term calculations break down.

To take a specific example, go back to the formula:

$$r = \rho + \theta \bar{g} - \frac{1}{2}\theta^2\sigma^2$$

which is based on the modelling assumption that $g\tilde{}\mathcal{N}\left(\bar{g}, \sigma^2\right)$. Weitzman'point is that, even if we agree with that specification, we know neither $\bar{g}$ nor σ. We will have to estimate them, and for this we need not only the data but an a priori distribution.

A standard way (Jeffreys prior) to choose such a distribution is to suppose that $\ln \sigma$ is Gaussian. If there are N experimental values available, we are led to a classical problem in statistics (find the variance of a Gaussian variable given N experimental values), the answer to which is a Student distribution with N degrees of freedom. It is well known that this distribution has fat tails. More precisely, if we have observed we find that:

$$\frac{1}{u'\left(c\left(0\right)\right)} E\left[u'\left(c\left(t\right)\right) \mid c\left(t_1\right), ..., c\left(t_N\right)\right] = +\infty$$

with the specification $u\left(c\right) = c^{1-\theta} / \left(1-\theta\right)$. In other words, given a finite number of observations, society should be willing to give up an unlimited amount of consumption today to gain any certain amount of consumption in the future. This corresponds to an interest rate of $r = -\infty$

3.5 Modifications 4: Equity and redistribution

3.5.1 The problems

Consider again the standard model: there is an infinite-lived representative consumer, who strives to maximize

$$\max \int_0^\infty u\left(C\left(t\right)\right) e^{-\rho t} dt \qquad \text{(Ramsey Growth Model)}$$

Problem 1: there is no such thing as a representative consumer People are different - in their tastes (utility function u), in their expectations (probability p). More importantly, some are rich, but most are poor. The first question is dealt with by aggregation theory (see the lectures by Jouini in this summer school). The second question, to my knowledge, has not attracted academic attention - except from Ramsey himself ! He devotes the last section of his seminal paper (1928) to this problem and concludes : "In such a case, therefore, equilibrium would be attained by a division of society into two classes, the thrifty enjoying bliss and the improvident at the subsistence level". It would of course be politically quite incorrect to mention the poor nowadays, and this is why academics gladly adhere to the fiction of the representative consumer.

Problem 2: no one lives for ever This means that the coefficient ρ (pure rate of time preference) will apply to different persons in the short to middle term (where the present generation is alive) and in the middle to long term (when we are all dead, and our descendants rule or are ruled). This means that this parameter is put to two different uses:

- for weighing consequences to me of my own actions
- for weighing consequences to others of my own actions

In a seminal paper, Sumaila and Walters (2005, [32]) separate the (psychological) impatience from the (ethical) concern for future generations. Their model combines three parameters:

- the population is renewed at the rate γ
- each generation has a pure rate of time preference ρ
- each generation discounts at the rate $\delta < \rho$ the utility of future generations

For an event which is to happen at time t, we find that the discount factor to apply is:

$$R(t) = e^{-\rho t} + \gamma \int_0^t e^{-\delta s} e^{-\rho(t-s)} ds \tag{23}$$

$$= (1 - \lambda) e^{-\rho t} + \lambda e^{-\delta t} \tag{24}$$

with:

$$\lambda = \frac{\gamma}{\rho - \delta} \tag{25}$$

. Note that this corresponds to a non-constant rate of time preference:

$$r(t) = -\ln\left((1 - \lambda) e^{-\rho t} + \lambda e^{-\delta t}\right)$$

- $r(t) \simeq -\delta$ in the long term
- $r(t) \simeq \lambda\rho + (1 - \lambda)\delta$ in the short term

4 Non-constant discount rates

4.1 Time inconsistency

Let us summarize the arguments for non-constant interest rates:

- there is strong experimental evidence that the psychological discount rate is not exponential, but hyperbolic: it is more like $(1+\mu t)^{-\alpha t}$ than $\exp(-\rho t)$
- even within the framework of the standard model, where the representative consumer has a constant psychological discount rate ρ, he/she ends up facing non-constant interest rates (see problem (13)), unless his/her utility happens to be CRRA. So, in the case of general utility functions, there will be time inconsistency even if the psychological discount rate is constant
- aggregating the beliefs of several individuals will lead to non-constant interest rates, as Weitzman pointed out, even if each of them has a constant psychological discount rate
- if the planner takes the interests of future generations into account, he/she will have to discount future welfare at a non-constant rate, even if he/she and the representative consumer both have constant rates of time preference, and even if this rate is the same !

My view is that this the case is overwhelming. Even if it were not, there is a final argument to be made: any policy recommendation drawn from the analysis must be robust to small changes in the model: one should ask, for instance, what becomes of the Ramsey model if the psychological discount rate is not exactly exponential. One would hope that there still is an optimal strategy, which converges to a stationary state, given by a suitably modified golden rule.

Unfortunately, non-constant interest rates create a special situation, known as *time inconsistency*. This is best explained in the framework of intergenerational equity, as explained at the end of the preceding section. The present generation faces the problem:

$$\max \int_0^\infty R(t)\, u(c(t))\, dt,$$

$$\frac{dk}{dt} = f(k(t)) - (n+g)\,k - c(t) \text{ and } k(0) = k_0$$

However, the present generation will not be around to implement the policy it has designed today, and must rely upon others (namely future generations) to do so. But the future generations may not agree with decisions taken on their behalf many years before they were around, and decide to carry out different ones. This is the *non-commitment* problem, which can be avoided only if the policy which seems optimal to us today still seems optimal to them when they are in charge. Unfortunately, with a non-constant rate of time preference, this will not happen.

Take two scenarios $c_1(\cdot)$ and $c_2(\cdot)$, both of which kick in at time T. In other words, $c_1(\cdot)$ and $c_2(\cdot)$ are defined for $s \geq T$. Say that we compare them, at some time $t_1 < T$, and we find $c_1(\cdot)$ is superior to $c_2(\cdot)$:

$$\int_T^{\infty} R(t-t_1)u(c_1(t))\,dt \geq \int_T^{\infty} R(t-t_1)u(c_2(t))\,dt \tag{26}$$

Let some time elapse, and do the comparison again at some later instant $t_2 < T$. Is it still true that we will find $c_1(\cdot)$ superior to $c_2(\cdot)$? This would mean that:

$$\int_T^{\infty} R(t-t_2)u(c_1(t))\,dt \geq \int_T^{\infty} R(t-t_2)u(c_2(t))\,dt \tag{27}$$

In the case when the discount rate is constant, so that $R(t) = \exp^{rt}$, the first inequality implies the second because of the special properties of the exponential function. We have:

$$\begin{aligned} \int_T^{\infty} R(t-t_2)u(c_1(t))\,dt &= \int_T^{\infty} e^{r(t-t_2)}u(c_1(t))\,dt \\ &= e^{r(t_1-t_2)}\int_T^{\infty} e^{r(t-t_1)}u(c_1(t))\,dt \\ &= e^{r(t_1-t_2)}\int_T^{\infty} R(t-t_1)u(c_1(t))\,dt \end{aligned}$$

so that (27) is derived from (26) by multiplying both sides by a constant.

In the case of non-constant discount rates, (26) no longer implies (27) ! In fact, a policy which is optimal for the decision-maker at time t_1, no longer is optimal for the decision-maker at a later time t_2 (even though the utility function $u(c)$ is unchanged). There is no control that will be simultaneously optimal for all those who will have to implement it.

The fact that this problem does not occur with constant interest rates is a miracle, which was already pointed out by Samuelson in 1937, with his usual foresight [31]: "our equations hold only for an individual who is deciding at the beginning of the period how he will allocate his expenditures over the period. Actually, however, as the individual moves along in time there is a sort of perspective phenomenon in that his view of the future in relation to his instantaneous time position remains time invariant, rather than his evaluation of any particular year (e.g. 1940). This relativity effect is expressed in the behaviour of men who make irrevocable trusts, in the taking out of life insurance as a compulsory savings measure, etc. The particular results we have reached are not subject to criticism on this score , having been carefully selected so as to take care of this provision. Contemplation of our particular equations will reveal that the results are unchanged even if the individual discounts from the existing point of time rather than from the beginning of the period." The last sentence, of course, means that Samuelson uses constant interest rates.

So, if interest rates are not constant, and there is no commitment technology, we will have to discard optimal control theory and build a new theory to replace it. This is a heavy mathematical work, which was carried out by Ekeland and Lazrak in ([10], and [11]). In the sequel, we describe their results

4.2 Equilibrium strategies

As above, we shall consider a general discount function $R : [0, \infty] \rightarrow R$. Throughout, it will be assumed to be continuously differentiable, with:

$$R(0) = 1,\ R(t) \geq 0,\ \int_0^\infty R(t)\,dt < \infty$$

and we consider the intertemporal decision problem (as it is seen at time $t = 0$)

$$\begin{array}{c} \max \int_0^\infty R(t)\,u(c(t))\,dt, \\ \frac{dk}{dt} = f(k(t)) - (n+g)\,k(t) - c(t) \text{ and } k(0) = k_0 \end{array} \tag{28}$$

Without loss of generality we are assuming that $g = n = 0$ (just change the definition of the production function $f(k)$)

Because of time-inconsistency, problem (28) can no longer be seen as an optimization problem.There is no way for the decision-maker at time 0 to achieve what is, from her point of view, the first-best solution of the problem, and she must turn to a second-best policy: the best she can do is to guess what her successors are planning to do, and to plan her own consumption $c(0)$ accordingly. In other words, we will be looking for a subgame-perfect equilibrium of a certain game.

A second idea now comes into play: we will assume perfect competition between decision-makers: none of them is sufficiently powerful to influence the global outcome. In the spirit of Aumann (see [3]) we will consider that the set of decision-makers is the interval $[0,\ T]$. At time t, there is a decision-maker who decides what current consumption $c(t)$ shall be. As is readily seen from the equation of motion $dk/dt = f(k) - c$, changing the value of c at just one point in time will not affect the trajectory. However, the decision-maker at time t is allowed to form a coalition with her immediate successors, that is with all $s \in [t,\ t+\varepsilon]$, and we will derive the definition of an equilibrium strategy by letting $\varepsilon \rightarrow 0$. In fact, we are assuming that the decision-maker t can commit her immediate successors (but not, as we said before, her more distant ones), but that the commitment span is vanishingly small.

We restrict our analysis to *Markov* strategies, in the sense that the policy depends only on a payoff relevant variable, the current capital stock and not on past history, current time or some extraneous factors. Such a strategy is given by $c = \sigma(k)$, where $\sigma : R \rightarrow R$ is a continuously differentiable function. If we apply the strategy σ, the dynamics of capital accumulation from $t = 0$ are given by:

$$\frac{dk}{ds} = f(k(s)) - \sigma(k(s)),\ k(0) = k_0$$

We shall say σ *converges to* $\bar{k}$, *a steady state* of σ, if $k(s) \longrightarrow \bar{k}$ when $s \longrightarrow \infty$, when the initial value k_0 is sufficiently close to $\bar{k}$. A strategy σ is *convergent* if there is some $\bar{k}$ such that σ converges to $\bar{k}$. In that case, the integral is obviously convergent, and its successive derivatives can be computed by differentiating under the integral. Note that if σ converges to $\bar{k}$, then we must have $f\left(\bar{k}\right) = \sigma\left(\bar{k}\right)$.

Suppose a convergent Markov strategy $c = \sigma(k)$, where $\sigma : R \to R$ is a continuously differentiable function, has been announced and is public knowledge. The decision maker begins at time $t = 0$ with capital stock k. If all future decision-makers apply the strategy σ, the resulting capital stock k_0 future path obeys

$$\frac{dk_0}{dt} = f\left(k_0(t)\right) - \sigma\left(k_0(t)\right), \quad t \geq 0 \tag{29}$$

$$k_0(0) = k. \tag{30}$$

We suppose the decision-maker at time 0 can commit all the decision-makers in $[0,\ \varepsilon]$,where $\varepsilon > 0$. She expects all later ones to apply the strategy σ, and she asks herself if it is in her own interest to apply the same strategy, that is, to consume $\sigma(k)$. If she commits to another bundle, c say, the immediate utility flow during $[0,\ \varepsilon]$ is $u(c)\,\varepsilon$. At time ε, the resulting capital will be $k + (f(k) - c)\,\varepsilon$, and from then on, the strategy σ will be applied which results in a capital stock k_c satisfying

$$\frac{dk_c}{dt} = f\left(k_c(t)\right) - \sigma\left(k_c(t)\right), \quad t \geq \varepsilon \tag{31}$$

$$k_c(\varepsilon) = k + (f(k) - c)\,\varepsilon. \tag{32}$$

The capital stock k_c can be written as $k_c(t) = k_0(t) + k_1(t)\,\varepsilon$ where:

$$\frac{dk_1}{dt} = \left(f'\left(k_0(t)\right) - \sigma'\left(k_0(t)\right)\right) k_1(t)\,, \ t \geq \varepsilon \tag{33}$$

$$k_1(\varepsilon) = \sigma(k) - c \tag{34}$$

and f' and σ' stand for the derivatives of f and σ. Summing up, we find that the total gain for the decision-maker at time 0 from consuming bundle c during the interval of length ε when she can commit, is

$$u(c)\,\varepsilon + \int_{\varepsilon}^{\infty} h(s)\, u\left(\sigma\left(k_0(t) + \varepsilon k_1(t)\right)\right) dt,$$

and in the limit, when $\varepsilon \to 0$, and the commitment span of the decision-maker vanishes, expanding this expression to the first order leaves us with two terms

$$\int_0^{\infty} h(t)\, u\left(\sigma\left(k_0(t)\right)\right) dt$$
$$+\,\varepsilon\left[u(c) - u(\sigma(k)) + \int_0^{\infty} h(t)\, u'\left(\sigma\left(k_0(t)\right)\right) \sigma'\left(k_0(t)\right) k_1(t)\, dt\right]. \tag{35}$$

where k_1 solves the linear equation

$$\frac{dk_1}{dt} = \left(f'\left(k_0(t)\right) - \sigma'\left(k_0(t)\right)\right) k_1\left(t\right), \quad t \geq 0 \tag{36}$$

$$k_1\left(0\right) = \sigma\left(k\right) - c. \tag{37}$$

Note that the first term of (35) does not depend on the decision taken at time 0, but the second one does. This is the one that the decision-maker at time 0 will try to maximize. In other words, given that a strategy σ has been announced and that the current state is k, the decision-maker at time 0 faces the optimization problem:

$$\max_c P_1\left(k, \sigma, c\right) \tag{38}$$

where

$$P_1\left(k, \sigma, c\right) = u\left(c\right) - u\left(\sigma(k)\right) + \int_0^\infty h\left(t\right) u'\left(\sigma\left(k_0\left(t\right)\right)\right) \sigma'\left(k_0\left(t\right)\right) k_1\left(t\right) dt. \tag{39}$$

In the above expression, $k_0\left(t\right)$ solves the Cauchy problem (29),(30) and $k_1\left(t\right)$ solves the linear equation (36),(37).

Definition 6 *A convergent Markov strategy $\sigma : R \to R$ is an* equilibrium strategy *for the intertemporal decision problem (28) if, for every $k \in R$, the maximum in problem (38) is attained for $c = \sigma\left(k\right)$:*

$$\sigma\left(k\right) = \arg\max_c P_1\left(k, \sigma, c\right) \tag{40}$$

The intuition behind this definition is simple. Each decision-maker can commit only for a small time ε, so he can only hope to exert a very small influence on the final outcome. In fact, if the decision-maker at time 0 plays c when he/she is called to bat, while all the others are applying the strategy σ, the end payoff for him/her will be of the form

$$P_0\left(k, \sigma\right) + \varepsilon P_1\left(k, \sigma, c\right)$$

where the first term of the right hand side does not depend on c. In the absence of commitment, the decision-maker at time 0 will choose whichever c maximizes the second term $\varepsilon P_1\left(k, \sigma, c\right)$. Saying that σ is an equilibrium strategy means that the decision maker at time 0 will choose $c = \sigma\left(k\right)$. Given the stationarity of the problem, if the strategy $c = \sigma\left(k\right)$ is chosen at time 0, it will be chosen at any future time t and as a result, the strategy σ can be implemented in the absence of commitment. Conversely, if a strategy σ for the intertemporal decision model (28) is not an equilibrium strategy, then it cannot be implemented unless the decision-maker at time 0 has some way to commit his successors.

4.3 The quasi-exponential case

From now on, we shall use the following specifications:

$$R\left(t\right) = \lambda \exp\left(-\delta t\right) + \left(1 - \lambda\right) \exp\left(-\rho t\right) \quad \text{and} \quad u\left(c\right) = \ln c \tag{41}$$

Using a logarithmic utility simplifies the computations, but the results extend to general CARA utilities $u(c) = c^{1-\theta}/(1-\theta)$, with $\theta > 0$, and presumably to more general utilities as well. The argument in [11] relies heavily the fact that the discount function is quasi-exponential. It has been extended to the case:

$$R(t) = (\lambda t + 1) e^{-\rho t}$$

by [39], and it would presumably extend to discount functions of the form $R(t) = \sum_{i=1}^{n} P_i(t) \exp(-\rho_i t)$, where the P_i are polynomials, although this has not been done.

Under the specifications (41), Ekeland and Lazrak have obtained an extension of the behaviour observed in the classical Ramsey model, with constant discount rate: they have found equilibrium strategies which converge to some asymptotic growth rate k_∞, independent of the initial capital k_0. The precise value of k_∞ must lie in an interval, which converges to the golden rule (3) when the discount function becomes exponential.

Denote by $\mathcal{K}$ is the *flow* associated with the differential equation (29) defined by

$$\frac{\partial \mathcal{K}(\sigma; t, k)}{\partial t} = f(\mathcal{K}(\sigma; t, k)) - \sigma(\mathcal{K}(\sigma; t, k))$$
$$\mathcal{K}(\sigma; 0, k) = k.$$

Definition 7 *Take some $k_\infty > 0$. We shall say that σ is a local equilibrium strategy converging to k_∞ if there is some open interval Ω around k_∞ such that σ is defined and C^2 on Ω, the flow $\mathcal{K}(\sigma; t, k)$ sends Ω into itself, and:*

- $\sigma(k) = \arg\max_c P_1(k, \sigma, c)$ *for all* $k \in \Omega$
- $\mathcal{K}(\sigma; t, k) \longrightarrow k_\infty$ *when* $t \longrightarrow \infty$, *for all* $k \in \Omega$

Here $P_1(k, \sigma, c)$ is given by formula (39).

Our main result then is:

Theorem 8 *Assume f is C^3 for $k > 0$. Define $\underline{k} \leq \bar{k}$ by:*

$$f'(\underline{k}) = \lambda\delta + (1-\lambda)\rho, \quad f'(\bar{k}) = \frac{1}{\frac{\lambda}{\delta} + \frac{1-\lambda}{\rho}} \tag{42}$$

Then, for every $k_\infty \in [\underline{k}\ ,\ \bar{k}]$, there exists a local equilibrium strategy converging to k_∞.

In the Ramsey case, when $\delta = \rho$, or $\lambda = 1$, we find the classical relation $f'(k_\infty) = \rho$. In the general case, the golden rule (3) is replaced by the inequality $\underline{k} \leq k_\infty \leq \bar{k}$ we find a continuum of possible equilibrium strategies, and corresponding asymptotic growth rates, and their range is fully characterized. So the proof in [11] is in two parts: first showing that every possible k_∞ is in that range, and then showing that every point in that range is a possible k_∞.

Note that there is still an indeterminacy, smaller of course as one nears the exponential case, but present. This indeterminacy arises from the fact that there are no boundary condition as $t \longrightarrow \infty$, nothing to replace the transversality condition at infinity of the exponential case. A further game-theoretical argument, however, will enable us to do away with that indeterminacy, and give a definite recommendation to the policy-maker.

Definition 9 *Let σ and σ' be two equilibrium strategies converging to k_∞ and k'_∞. We shall say that σ is* eventually Pareto-dominated by *σ' if, for any starting point k, there is some $t > 0$ such that:*

- *for all $s > t$, the decision-maker at time s prefers σ' to σ*
- *if one applies strategy σ' after time t, it remains true that at all subsequent times, the decision-makers prefer σ' to σ*

In [[11] it is proved that, whenever a strategy converges to some k_∞, it is eventually Pareto-dominated by any strategy that converges to some $k'_\infty < k_\infty$. So the only equilibrium strategy which is not Pareto-dominated is the leftmost none, namely the one which converges to $\underline{k}$. Let us express this result:

Proposition 10 *All convergent equilibrium strategies are Pareto-dominated, except the one(s) which converge(s) to $\underline{k}$*

So the rational choice is now clear: it is the equilibrium strategy which converges to the point $\underline{k}$ where

$$f'(\underline{k}) = \frac{1}{\left(\frac{\lambda}{\delta} + \frac{(1-\lambda)}{\rho}\right)} \tag{43}$$

Indeed, for any other choice, one of the future decision-makers will switch to another strategy, in the knowledge that his/her successors will follow suit. So there is no point in applying now a strategy which one knows will not be implemented later on, even if one has to wait for the distant future.

Applying the Sumaila-Walters specifications (23), (24) and (25) gives:

$$\begin{aligned} f'(\underline{k}) &= \frac{1}{\left(\frac{\gamma}{\delta(\rho-\delta)} + \frac{\rho-\delta-\gamma}{\rho(\rho-\delta)}\right)} = \frac{\rho\delta(\rho-\delta)}{\gamma\rho + \rho\delta - \delta\delta - \gamma\delta} \\ &= \rho\left(1 - \frac{1}{1+\delta/\gamma}\right) \end{aligned}$$

which is to be compared to the golden rule (3), where we we have taken logarithmic utility, so that $\theta = 1$, and $n = 0$ (no technological progress), so that it becomes $f'(k_\infty) = \rho$. So concerns for intergenerational equity *lower* the interest rate, by a factor which depends only on δ/γ. For instance, if $\delta = \gamma$, that is, if we discount the utility of future generations at a rate which is precisely equal to the renewal rate of the population, then the planner should replace ρ, the pure rate of time preference of the present generation, by $\rho/2$.

4.4 Bibliographical notes

The first paper to investigate the Ramsey model of economic growth with non-constant discount rates is due to Barro [5], who investigated the case of logarithmic utility. There was an earlier literature, in the discrete-time framework, originating with the seminal papers of Strotz [35] and Phelps and Pollack [27]. Going from the discrete to continuous time proved to be mathematically challenging. Ekeland and Lazrak ([10], and [11]) then introduced the idea of perfect competition between decision-makers, which enabled them to characterize seem to fully characterize equilibrium strategies in this case and to derive an analogue of the HJB equation.

5 Conclusion

This introduction to long-term interest rates, incomplete as it is, would be even more so if I failed to direct the reader to [[23]] and [[28]], which are standard references in the field. As a personal conclusion, I would like to remind the reader once more that determining the proper interest rates to use for projects with very long-term consequences, such as those which impact the environment, is one of the most important ways that the economic profession can contribute to solving the major challenges which our planet faces today. Such interest rates should incorporate the distributional and ethical concerns of our contemporaries. In other words, economics, after decades of riding the tiger of economic expansion, should once more become a normative science and take the lead.

References

[1] Aghion, Philippe and Howitt, Peter (1998) "*Endogeneous growth theory*", MIT Press

[2] Aubin, Jean-Pierre and Clarke, Frank (1979) "*Shadow prices and duality for a class of optimal control problems*", SIAM J. Control Optim. 17 , pp. 567–586.

[3] Aumann, Robert "*Markets with a continuum of traders*", Econometrica, 32 (1964), p. 39-50

[4] Barro, Robert and Sala-i-Martin, Xavier (1995) "*Economic growth*", McGraw and Hill

[5] Barro Robert J. “*Ramsey Meets Laibson in the Neoclassical Growth Model",* Quarterly Journal of Economics 114 (November 1999): 1125-52.

[6] Blanchard, Olivier and Fisher, Stanley (1994) "*Lectures on macroeconomics*", MIT Press

[7] Bressan, Alberto and Piccoli, Benedetto (2007) "I*ntroduction to the Mathematical Theory of Control*" Applied. Math Series vol. 2, American Institute of Mathematical Sciences

[8] Clarke, Frank, (1983), "Optimization and Nonsmooth Analysis", Wiley; reprinted (1990), SIAM, series "Classics in Applied Mathematics"

[9] Boehm-Bawerk, Eugen von (1884), "*History and Critique of Interest Theories*"

[10] Ekeland, Ivar and Lazrak, Ali (2005) "*Being serious about non-commitment*", http://arxiv.org/abs/math/0604264

[11] Ekeland, Ivar and Lazrak, Ali (2008) "*Equilibrium policies when preferences are time inconsistent*", http://arxiv.org/abs/0808.3790

[12] Ekeland, Ivar and Scheinkman, Jose (1986) "*Transversality conditions for some infinite horizon discrete time opimization problems*", Mathematics of Operations Research 11, p. 216-229

[13] Ellsberg, Daniel (1961), "*Risk, Ambiguity, and the Savage Axioms*", Quarterly Journal of Economics 75 (4), p. 643–66

[14] Fisher, Irving (1930) "*The Theory of Interest*"

[15] Gierlinger, Johannes and Gollier, Christian, "*Socially Efficient Discounting under Ambiguity Aversion*", IDEI Working Paper, n. 561, November 2008.

[16] Guesnerie, Roger (2004) "*Calcul economique et developpement durable*", Revue Economique, 55 (3), p. 363-382

[17] Guesnerie, Roger, Guéant, Olivier, and Lasry, Jean-Michel (2008), "*Economic Calculus and Sustainable Development*", Cahiers de la Chaire de Développement Durable, Université Paris-Dauphine http://www.ifd.dauphine.fr/fileadmin/mediatheque/recherche_et_valo/FDD/CAHIER16_GUESNERI

[18] Hammond and Phelps,

[19] Heal, Geoffrey (2007) "*Climate change economics: a meta-review and some suggestions*"

[20] Kamihigashi, Takashi (2001) "*Necessity of Transversality Conditions for Infinite Horizon Problems*". Econometrica 69 (4), 995-1012.

[21] Klibanoff, P., Marinacci, M., and Mukerji, S. (2005) "*A smooth model of decision making under ambiguity*" Econometrica 73, p. 1849-1892

[22] Koopmans, Tjalling (1960), "*Stationary utility and impatience*", Econometrica 28 (2), p. 287-290

[23] Lind, Robert C., editor (1982), "*Discounting for Time and Risk in Energy Policy*", Resources for the Future, Washington D.C, distributed by John Hopkins Press

[24] Michel, Philippe (1982) "*On the Transversality Condition in Infinite Horizon Optimal Problems.*" Econometrica 50, p. 975-985.

[25] Nordhaus, William D. (1991) "*To slow or not to slow: the economics of the greenhouse effect*" Economic Journal 101 p. 920-937

[26] Nordhaus, William D. (1992) "*Lethal model 2: the limits to growth revisited*" Brookings papers on economic activity (2) P. 1-43

[27] Phelps, Edmund S., and Pollak, Robert A. "*On Second-Best National Saving and Game-Equilibrium Growth.*' *Review of Economic Studies* 35 (April 1968): 185-99.

[28] Portney, Paul and Weyant, John (1999), "Discounting and Intergenerational Equity", Resources for the Future, Washington D.C.

[29] Rae, John (1834) "*Statement of some new Principles on the subject of Political Economy*"

[30] Ramsey, Frank (1928) "*A mathematical theory of savings*", The Economic Journal, 38 (152), p 543-559

[31] Samuelson, Paul (1937) ""*A note on measurement of utility*", Review of Economic Studies 4, p. 155-61

[32] Sumaila, Ussif and Walters, Carl (2005) "I*ntegenerational discounting: a new intuitive approach*" Ecological Economics 52 p. 135-142

[33] Ramsey, Frank (1928) "A mathematical theory of saving", Economic Journal 38, p. 543-559

[34] Romer, David (1996), "*Advanced macroeconomics*", McGraw and Hill

[35] Strotz, Robert H. "*Myopia and Inconsistency in Dynamic Utility Maximization*" *Review of Economic Studies* 23, no. 3 (1956): 165-80.

[36] Weitzman, Martin (2001) "*Gamma discounting*", American Economic Review 91 (1) p. 260-271

[37] Weitzman, Martin (2008) "*On modeling and interpreting the economics of catastrophic climate change*", American Economic Review 91 (1) p. 260-271

[38] Weitzman, Martin (2007) "*A review of the Stern review on the economics of climate change*", Journal of Economic Literature vol. 45, no3, pp. 703-724

[39] Qinglong Zhou, personal communication, Nankai University, June 23, 2009